THE GENETIC ARCHITECTURE OF MORPHOLOGICAL CHANGES IN THE NINESPINE STICKLEBACK (*PUNGITIUS PUNGITIUS*) AND THE DOMESTICATED PIGEON (*COLUMBA LIVIA*): INSIGHTS FROM POPULATION ANALYSIS, QUANTITATIVE TRAIT MAPPING AND WHOLE-GENOME RESEQUENCING

by

Sydney Ann Stringham

A dissertation submitted to the faculty of
The University of Utah
in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

Department of Biology

The University of Utah

August 2014

The University of Utah Graduate School

STATEMENT OF DISSERTATION APPROVAL

The dissertation of **Sydney Ann Stringham**

has been approved by the following supervisory committee members:

Michael D. Shapiro	, Chair	**4/30/2014** Date Approved
Richard M. Clark	, Member	**4/30/2014** Date Approved
Jon Seger	, Member	**4/30/2014** Date Approved
Gabrielle Kardon	, Member	**4/30/2014** Date Approved
Lynn B. Jorde	, Member	**4/30/2014** Date Approved

and by **Neil Vickers**, Chair of

the Department of **Biology**

and by David B. Kieda, Dean of The Graduate School.

ABSTRACT

Here I present the results of my doctoral dissertation, which is aimed at increasing our understanding of the genetic basis of large morphological changes. The work I have done has primarily been carried out using ninespine sticklebacks (*Pungitius pungitius*) as a model organism. Specifically, I have investigated the genetic architecture of pelvic reduction (a structure homologous to tetrapod hindlimbs) in multiple populations using a combination of traditional QTL mapping as well as whole-genome comparisons. Additionally, I examined the structure of breed groups among domesticated pigeons (*Columba livia*) in order to determine whether or not similar derived traits are found in genetically unrelated breeds. This work lays the foundation to develop the domesticated pigeon as a genetic and developmental model for avian diversity.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

CHAPTER 1

COMPLEX GENETIC ARCHITECTURE UNDERLIES PELVIC REDUCTION IN A CANADIAN POPULATION OF NINESPINE STICKLEBACK (*PUNGITIUS PUNGITIUS*): A COMBINED GENOMIC RESEQUENCING AND QTL MAPPING APPROACH

Introduction

Although there are many examples of divergent vertebrate lineages evolving similar traits, relatively little is known about the types and number of mutations underlying these convergent events. Stickleback fish provide an ideal model to investigate dramatic convergent events because they are one of a few groups that have evolved anatomical, physiological, or behavioral differences among populations of the same species that are of a magnitude typically seen between different species. This allows powerful methods such as quantitative trait locus (QTL) mapping used to investigate the evolution of dramatic, and adaptively relevant, changes in wild populations. The retreat of glacial ice less than 20,000 years ago allowed populations of marine sticklebacks to colonize new, inland freshwater habitats (Bernatchez and Wilson 1998; Hewitt 2000). This shift to freshwater presented novel trophic niches as well as new physiological and predatory challenges, and many geographically and phylogenetically distinct populations of sticklebacks evolved to their new habitats in similar ways.

One dramatic example of a repeated change is the loss of the pelvic complex (Bell and Foster 1994). The stickleback pelvis is homologous to the tetrapod hindlimb and is composed of a pelvic girdle and two serrated spines. While the pelvis provides protection from gape-limited predators (Hoogland et al. 1957; Hagen and Gilbertson 1972; Moodie 1972; Gross 1978; Lescak and von Hippel 2011), it is thought to be disadvantageous when grasping predators are a bigger threat (Hoogland et al. 1957; Reimchen 1980; Reist 1980; Bell et al. 1993; Bell and Orti 1994; Ziuganov and Zotin 1995; Marchinko 2009). Interestingly, parallel reduction in the pelvic skeleton has occurred not only among multiple populations of threespine sticklebacks (*Gasterosteus aculeatus,* which has been the subject of many classic behavioral, ecological, and recent genomic studies), but also in the ninespine stickleback (*Pungitius pungitius*) and the brook stickleback (*Culaea inconstans*) (Nelson and Atton 1971; Wootton 1976; Blouw and Boyd 1992; Bell and Foster 1994; Ziuganov and Zotin 1995). Therefore, the stickleback family (Gasterosteidae) is an ideal multispecies system to examine the genetics of adaptive traits on both micro- and macroevolutionary scales.

Several studies have examined the genetic basis of pelvic reduction in threespine stickleback populations. First, mapping studies showed that a major-effect quantitative trait locus (QTL) on linkage group 7 along with between three and four minor-effect QTL (Cresko et al. 2004; Shapiro et al. 2004; Coyle et al. 2007; Shapiro et al. 2009; Shikano et al. 2013). This region of the genome was interesting because it contained the hindlimb specific transcription factor *Pitx1* (Shapiro et al. 2004). Later, Chan et al. (2010) confirmed that independent deletions of a pelvic enhancer of *Pitx1* were associated with pelvic reduction in several populations. Thus, the same phenotype in different

populations is controlled by independent mutations in the same gene.

Complementation tests (Shapiro et al. 2006a) and QTL mapping (Shikano et al. 2013) also identified *Pitx1* as a candidate for pelvic reduction in ninespine sticklebacks. Therefore, the same gene might be responsible for similar morphological changes in two species that have been separated by more than 10 million years. However, unlike in threespine sticklebacks, pelvic reduction in ninespine sticklebacks does not map to *Pitx1* in all populations examined to date. For example, a study of an Alaskan population found that the major contributor to pelvic reduction mapped to linkage group 4, which is unlinked to *Pitx1* (Shapiro et al. 2009). Therefore, at least two different genetic changes potentially lead to pelvic reduction in ninespine sticklebacks.

With only a handful of examples, it is difficult to make broad conclusions about the genetic architecture of pelvic reduction in ninespine sticklebacks. Recent phylogenetic analysis suggests that, in North America, ninespine sticklebacks probably dispersed from three distinct refugia and evolved pelvic reduction independently in each case. Populations from the west coast of North America descended from populations from the Bering refugium, inland populations dispersed from the Mississippi refugium, and those found along the east coast of North America came from the Atlantic refugium. Given that the two mapping studies conducted thus far in ninespine sticklebacks come from the Bering lineage in North America (Shapiro et al. 2009) or an Eastern European lineage (Shikano et al. 2013); a closer investigation of the genetic mechanisms underlying pelvic reduction in a population derived from the Mississippi refugium could provide insight into general genomic patterns of pelvic reduction in ninespine sticklebacks.

To this end, we examined two populations of ninespine sticklebacks from the Northwest Territories of Canada, Salt River and Pine Lake, which exhibit an unusually broad range of pelvic phenotypes. That is, while most freshwater populations of ninespine sticklebacks have a complete pelvic skeleton and a few exhibit pelvic loss in all individuals, these populations comprise individuals with a wide range of pelvic phenotypes.

In this study, we took a two-step approach, combining traditional QTL mapping and comparative whole-genome sequencing, to identify genomic regions that contribute to pelvic phenotype in Salt River and Pine Lake. We began by conducting traditional QTL mapping in the Salt River population. QTL mapping is a robust method to identify genomic regions that contribute to phenotypic variation, and it has relatively low rates of false positives (Sahana et al. 2006). However, in laboratory crosses with limited numbers of progeny, QTL mapping can often result in candidate genomic regions that contain hundreds of genes. To address this challenge, we also assembled a draft genome for the ninespine stickleback and used it as the basis for whole-genome resequencing aimed at identifying genomic regions with divergent allele frequencies, and presumably selection, between phenotypic classes. In contrast to linkage mapping in laboratory populations, whole-genome sequencing studies of trait variation in natural populations can potentially implicate smaller genomic regions. This precision results from historical recombination events that break down linkage disequilibrium between causative mutations and their surrounding (presumably neutral) variants. However, the large datasets in such studies are more prone to false positives. A combination of these approaches should allow for the identification of QTL with relative confidence, which can then be validated and narrowed

with resequencing data (Stinchcombe and Hoekstra 2008).

Materials and Methods

Genome Sequencing and Assembly

The DNA for reference genome sequencing was extracted from a single female fish from an unnamed creek in Wasilla, Alaska (61° 37' N, 149° 30' W). This population was chosen because it has low rates of heterozygosity compared to other populations (Aldenhoven et al. 2010), which facilitates genome assembly (Holt et al. 2002; Vinson et al. 2005). We constructed two paired-end sequencing libraries with insert sizes of 250 bp and 500 bp using the Illumina Paired-End DNA Sample Prep Kit. An additional mate-pair library with an insert size of 2400 bp was also constructed (National Center for Genome Resources, Santa Fe, New Mexico). 101-bp paired-end sequencing was performed on all libraries using the Illumina HiSeq2000 platform (University of Utah High Throughput Genomics core). Statistics of the raw reads are listed in Table 1.1.

An initial genome assembly was constructed using ALLPATHS-LG (r40776) (Gnerre et al., 2011) and contained 8784 scaffolds with a total length of 387.6 MB. We improved the assembly by using SSPACE (Boetzer et al., 2011) to perform further scaffolding with end reads from two ninespine stickleback BAC libraries (92,160 reads, mean length 910 bp; BAC libraries VMRC34 and VMRC35, Benaroya Research Institute at Virginia Mason, Seattle, WA). Because SSPACE uses short paired-end reads as input, we used the following protocol to convert the Sanger reads into a short-read library: for each fragment in a read pair, we split the read into 2 segments and selected the first 80 bp from each segment to construct a new paired-end library in silico. By aligning this library to the previous genome assembly with Bowtie (Langmead et al., 2009), we calculated a

mean BAC insert size of 140 kb (SD = 41 kb). We then ran SSPACE to scaffold our genome assembly with the new library.

The final genome assembly contains 7824 scaffolds with a total length of 428.1 Mb and contig and scaffold N50 lengths of 122.8 kb and 302.8 kb, respectively. The longest contig is 1.16 Mb and the longest scaffold is 4.14 Mb. Expected genome size based on kmer distribution in the 250- and 500-bp libraries calculated using Jellyfish (Marcais and Kingsford 2011) was 518.4 Mb. Although this longer than our assembled genome size, the CEGMA (Parra et al. 2007) pipeline reports that 90.32% of conserved eukaryotic proteins were found within this assembly, indicating a relatively complete gene annotation and suggesting that unassembled genomic regions are probably enriched for repetitive sequences.

Finally, to place scaffolds in a relative genomic order we identified regions of synteny between threespine (BROADS1 assembly) and ninespine stickleback genomes. Using protein sequence, we identified reciprocal best BLAST hits between the two genomes. Any ninespine scaffold that had at least one reciprocal best blast hit on a threespine chromosome was considered orthologous and was placed in a relative order based on the threespine stickleback genome. Fifteen ninespine stickleback scaffolds showed synteny with two different threespine stickleback chromosomes, these scaffolds were not assigned a relative position. A total of 2,672 ninespine stickleback scaffold (83% of the genome assembly) showed synteny with a single threespine stickleback chromosome.

Annotation

MAKER version 2.29 (Holt and Yandell 2011) was used to annotate the genome assembly using multiple lines of evidence. An RNA-seq library was created using mRNA extracted from adult heart, eye, brain, liver, and muscle tissue as well as whole embryos at 3 and 6 days postfertilization (chorion and yolk removed). mRNA samples from all tissues were combined in equimolar amounts for Illumina library construction. RNA-seq reads were assembled using Trinity (Grabherr et al. 2011) and provided as evidence for the genefinders in MAKER. Additional evidence included all RefSeq teleost proteins (downloaded July 30, 2013 from http://www.ncbi.nlm.nih.gov) and all Uniprot/SwissProt proteins (downloaded July 29, 2013 from ftp://ftp.uniprot.org/pub/databases/uniprot/current_release/knowledgebase/complete).

Repetitive regions were masked using a species-specific repeat library generated by RepeatModeler (Smit and Hubley 2008). This library was then aligned to Uniprot/SwissProt proteins using BLAST (Altschul et al. 1990) ($E < 0.0001$) and repeat library entries with matches to a known protein gene were removed. Additional masking was done with a list of known transposable elements provided by MAKER. Other areas of low complexity were soft-masked (Korf et al. 2003) using Repeatmasker (Smit et al. 1996) to prevent the seeding of evidence alignments in those regions but still allowing extension of evidence alignments through them (Altschul et al. 1990; Cantarel et al. 2008). Genes were predicted using SNAP (Korf 2004) and Augustus (Stanke and Waack 2003; Stanke et al. 2008) trained for *Pungtius pungitius* using MAKER in an iterative fashion (Cantarel et al. 2008).

The final annotation set consisted of the all MAKER-generated annotations with

protein or mRNA-seq support and the subset of the unsupported gene predictions that contained one or more protein family domains as detected by IPRscan (Quevillon et al. 2005). In total, we identified 22,432 protein-coding genes (mean length = 10,015 bp). Of these, 21,516 showed homology with other species, and 16,654 were supported by mRNAseq data. Overall, the ninespine stickleback genome is similar to other published teleost fish genomes in terms of the number of annotated genes, but is relatively compact in length in comparison to several other teleosts (Table 1.2).

Mutation Rate in Ninespine Lineage

We estimated the species-specific mutation rate for the ninespine stickleback as described previously by Shapiro et al. (2013). Briefly, we used TBLASTX (Altschul et al. 1990) alignments ($E < 10^{-8}$) to identify one-to-one orthologs between fugu (*Takifugu rubripes*), threespine stickleback, and ninespine stickleback, with fugu aligned to threespine and ninespine stickleback separately. We then identified four-fold degenerate codon positions shared between the three species and generated three way alignments from 5,773 orthologous genes. We ran MODELTEST (Posada and Crandall 1998) using these alignments and found that the General Time Reversible (GTR) substitution model best fit the observed data. We then ran the baseml script in the PAML package (Yang 2007) under the GTR model with divergence times of 85 MYA and 100 MYA based on reports by Near et al. (2012) and Santini et al. (2009), respectively.

Population Structure Analysis

We assessed population structure within and between the Salt River and Pine Lake populations using the Bayesian clustering analysis in Structure with a 50,000-

iteration burn-in followed by 500,000 iterations (Pritchard et al. 2000). We genotyped 186 Pine Lake fish and 160 Salt River fish at 12 unlinked microsatellite loci: Pun44, Pun68, Pun157, Pun78, Pun117, Pun255, Pun212, Pun203, Pun171, Pun19, Pun261, and Pun238 (Shapiro et al. 2009). We examined population models from K = 2 to K = 6.

Pooled Resequencing

Fish with the most extreme pelvic phenotypes, that is, fish with a pelvic score of 8 (complete) and fish with a pelvic score of 0–2 (reduced) (Bell et al. (1987)), were collected from Salt River (complete, n = 100; reduced, n = 64) and Pine Lake (complete, n = 100; reduced, n = 89) (Northwest Territories, Canada; Pine Lake: 59° 33'N, 112° 15' W; Salt River: 59° 49'N, 111° 58' W). Equimolar amounts of DNA from each individual were combined into one of four pools of DNA (Pine Lake complete, Pine Lake reduced, Salt River complete, Salt River reduced). These pools were used to construct Illumina sequencing libraries with an insert size of 250 bp, which were sequenced using the Illumina HiSeq2000 platform to a depth of 25–45x coverage using 101-bp paired end reads (University of Utah High Throughput Genomics Core).

Variant Calling

We aligned sequencing reads from each population pool to the reference genome using Bowtie2 (Langmead and Salzberg 2012). We then used two different software pipelines to call nucleotide variants from the resulting BAM alignment files. First, SNVer (Wei et al. 2011), software specifically designed to detect variants in pooled sequences, identified 5.2 million variants. Second, the Genome Analysis Toolkit (Van der Auwera et al. 2013) was used to realign indels (RealignerTargetCreator, IndelRealigner) and call

variants with UnifiedGenotyper; this method identified 5.4 million variants. The two variant sets were then intersected to include only SNPs that were identified by both methods, resulting in a final variant set of 3,510,585 SNPs.

F_{ST} and Likelihood Ratio Test Analyses

A number of metrics have been used in whole-genome comparisons, but not all are applicable for pooled sequencing as many require individual haplotype information and are not designed to properly account for sequencing errors in pooled data. In order to assess allele frequency differentiation between phenotypic classes we used both F_{ST} (Weir and Cockerham 1984) and a likelihood ratio test (Kim et al. (2010). The latter method was included because it includes depth of coverage as a factor, which is important because regions of low coverage may not accurately reflect the allele frequency of the population. For both tests, we excluded sites with less than 10x coverage to avoid variants that might not accurately reflect allele frequencies in the population (Zhu et al. 2012). We also excluded sites with greater than 100x coverage, as these sites are probably repetitive sequences that do not map uniquely. Both F_{ST} and LRT metrics were smoothed over a 10-kb sliding window with 2 kb steps. Depending on the population and metric used, the number of total windows analyzed for each population and pool was between 174,914 and 175,169.

Cross Husbandry

Twenty-eight crosses were made between individuals from the Salt River population. Offspring were raised to at least 30-mm standard length in 29-gallon aquaria with a 16-hour light/8-hour dark cycle. Fish were euthanized using MS-222 and

preserved in 70% ethanol. Tissue was removed from the liver and right pectoral fin for subsequent DNA isolation. To stain external bone, fish were fixed in 10% neutral buffered formalin, stained with alizarin red, and preserved in 70% ethanol for phenotyping.

Four half-sibling families (12 crosses in total) showed variation in pelvic phenotype in the F1 generation and were used for QTL mapping (Table 1.3). Pelvic score in these crosses ranged from 0 (no pelvis) to 8 (complete pelvis). In total, 381 F1 offspring were included in subsequent analyses.

Phenotyping

For all crosses, skeletal measurements were taken using digital calipers under a dissecting microscope. Measurements included standard length, pelvic girdle length, pelvic spine length, and pelvic ascending process height (Shapiro et al. 2009). Separate measurements were taken on left and right sides. The same person measured individual traits, and each measurement was taken three times then averaged. Numbers of lateral plates were assessed separately for left and right sides and, because mid-body plates were absent in all individuals, we counted anterior and posterior rows separately.

Bulked Segregant Analysis of Crosses

Using crosses from Salt River established in 2010, DNA was extracted from individuals with extreme pelvic phenotypes (that is, those with a pelvic score of 8 or those with a pelvic score of 0–2) and pooled, within crosses, in equimolar amounts (see Table 1.4). These pools, along with individual (unpooled) parents, were genotyped with 192 microsatellite markers as previously described (Shapiro et al. 2009), and the results

were visualized with GeneMapper software (Applied Biosystems, Foster City, CA). The relative PCR amplification intensities of alleles from each pool were compared by eye, and we found that 32 markers showed differential allele representation between the two phenotypic pools in at least two of the three families tested.

These 32 markers were then used to genotype individual fish from four families from 2008 and 2010. Additionally, because the QTL mapping software we used to analyze our crosses is not able to process linkage groups with a single marker, Pun61, Pun98, Pun159, Stn259, Stn329, and Stn435 were added to anchor markers that were the only representative on a given linkage group. Since many progeny of the crosses had immature gonads, Stn19 was also used to genotype all fish individually in order to determine sex (Shikano et al. 2011). This marker produced a discernable genotype in 88% and 79% of fish in 2008 and 2010 crosses, respectively.

QTL Mapping

The number of offspring in each of the 12 crosses was between 15 and 49. By grouping crosses with the same male parent and conducting all further mapping analyses with F1 half-sibling families we were able use larger sample sizes and improve our power to identify quantitative trait loci (Family 1, n = 178; Family 2, n = 56; Family 3, n = 89; Family 4, n = 58). To that end, we used the half-sibling portal of GridQTL (Seaton 2006) to analyze each family separately. We included genotype data for 38 markers, phenotype data for 8 pelvic traits, standard length and sex, and marker distances from a previously published ninespine linkage map (Shapiro et al. 2009). GridQTL was run using length as a covariate, sex as a cofactor, and 1000 chromosome-wide permutations. Default settings were used for all other parameters.

Results and Discussion

Draft Genome and Comparative Resequencing

Genome sequencing and assembly. The ninespine stickleback reference assembly was sequenced from a single fish from the Church Road population (south-central Alaska) using the Illumina HiSeq2000 platform. The draft genome contains 7,850 scaffolds (N50 length = 299.5kb) and a total assembled length of 441.1 MB with approximately 140.5x coverage. See Table 1.5 for a summary of genome metrics.

The final annotation set ("MAKER standard build"; Campbell et al. 2014), contains 22,432 protein coding genes, 80.1% of which contain a protein domain as detected by IPRscan (Quevillon et al. 2005). 87% of genes have an annotation edit distance less than 0.5, consistent with a well annotated genome (Holt and Yandell 2011), and 95.9% of the annotated genes are similar to proteins in SwissProt as identified by BLAST ($E < 0.0001$) (Altschul et al. 1990).

We used annotated protein sequence from threespine stickleback and fugu (*Takifugu rubripes*) to calculate a lineage-specific mutation rate of between 0.009 and 0.010 mutations/site/MY assuming divergence times of 85 MY and 100 MY, respectively (Santini et al. 2009; Near et al. 2012). These values are similar to mutation rates calculated for other teleost fish lineages (0.007–0.04 mutations/site/MY) (Jaillon et al. 2004; Burridge et al. 2008).

Pooled resequencing. Pooled resequencing is an effective method for identifying genomic regions that are differentiated or under selection when single-genome resequencing is cost-prohibitive or impractical. The estimation of allele frequencies in pooled data has been shown to accurately represent the allele frequencies in pooled

populations given a minimum of 10x coverage (Zhu et al. 2012; Rellstab et al. 2013). This method has been used to successfully identify selective sweeps and candidate genomic regions underlying traits in many organisms, including *Arabidopsis,* maize, *Drosophila*, domestic chickens and pigs, and humans (Burke et al. 2010; Marklund and Carlborg 2010; Rubin et al. 2010; Turner et al. 2010; Janssen et al. 2011; Udpa et al. 2011; Zhou et al. 2011; Rubin et al. 2012).

The Salt River and Pine Lake populations (Figure 1.1 A) of ninespine sticklebacks present a unique opportunity to use this method to better understand the genetic architecture of pelvic reduction. The majority of freshwater ninespine stickleback populations have a complete pelvic skeleton and at least 10 exhibit a reduced pelvis; however, few show the range of pelvic phenotypes seen in Salt River and Pine Lake (Figure 1.1 B, C) (Nelson 1971; Nelson and Paetz 1972; Blouw and Boyd 1992; Ziuganov and Zotin 1995; Shapiro et al. 2006b; Herczeg et al. 2010; Mobley et al. 2011; Klepaker et al. 2013). Comparing very divergent phenotypes within a single population controls for demographic differences in genetic background that could confound results when comparing phenotypes between two populations.

We began by genotyping individuals that contributed to the resequencing pool with 12 unlinked microsatellite markers and confirmed that there was no genetic differentiation based on pelvic phenotype within these populations. We analyzed Pine Lake and Salt River samples together. Using the Evanno method (Evanno et al. 2005) as implemented in STRUCTURE HARVESTER (Earl and vonHoldt 2011) determined that the most likely number of populations is 2 (Figure 1.2); there was no substructure seen between phenotypes within a population.

We then collected and pooled DNA from between 68 and 100 fish with the most extreme pelvic phenotypes from each population (for a total of 4 groups) and sequenced each pool to a depth of between 25–45x coverage. Because pools were assembled based only on pelvic phenotype in unstructured populations, some of the genomic regions that show signatures of selection or differentiation are expected to influence pelvic variation.

Allele frequency differences between pools of distinct phenotypes within a population were assessed using both F_{ST} (Weir and Cockerham 1984) and a likelihood ratio test (LRT) (Kim et al. 2010) in 10kb sliding windows, with 2kb steps, along the length of the genome. The mean values for each test when comparing phenotypes within a population were as follows: Salt River, mean LRT = 0.512, mean F_{ST} = 0.016; Pine Lake, mean LRT = 0.362, mean F_{ST} = 0.007. The overall low value for both of these statistics is expected, as both are metrics of population differentiation and the phenotypic groups that are being compared originate from the same population.

To identify genomic regions that might contribute to pelvic phenotype in these populations, we further examined the top 0.1% of all windows (Salt River: $F_{ST} \geq 0.060$, LRT $\geq$ 2.00; Pine Lake, $F_{ST} \geq 0.022$, LRT $\geq$ 1.24). In Salt River, all linkage groups except 2, 10, and 21 have windows that are in the top 0.1%. In Pine Lake, all linkage groups have windows that meet this threshold. Additionally, similar regions of the genome show elevated F_{ST} and LRT values when comparing pelvic complete and pelvic reduced pools in Salt River and in Pine Lake (Figures 1.3 and 1.4). In both populations, linkage group 12 is enriched for elevated F_{ST} and LRT scores. There are also regions in the center of linkage group 4, the end of linkage group 19 and a segment of the unordered region of the genome that are elevated in both populations.

Based on the geographic proximity and similarity of pelvic phenotypes seen in both of these populations, we hypothesized that similar regions of the genome may be affecting pelvic phenotype in these populations. In order to test for overlap between the results from Salt River and those from Pine Lake, we counted how many of the windows found in the top 0.1% of LRT scores in Salt River overlap exactly with any windows in the top 0.1% identified in Pine Lake. Overall, we found that of the windows in the top 0.1% in both Salt River and Pine Lake, 259 overlapped exactly between the two populations (14.8% of all windows in the top 0.1%).

There are also regions of the genome with high differentiation between phenotypes in one population, but not the other. For example, linkage group 10 has a region of prominent differentiation in Pine Lake, but not in Salt River. Additionally, linkage group 21 contains 23 windows in the top 0.1% of scores in Pine Lake, while there are no high-LRT windows on that linkage group in Salt River.

Overall, these patterns suggest that some genomic regions are indeed associated with pelvic reduction in both populations, for example, linkage groups 4, 12, and 19. However, there are also other genomic regions associated with pelvic phenotype that are unique to one population (e.g., linkage group 10 in Pine Lake).

Quantitative Trait Mapping

Crosses and bulked segregant analysis. Because whole-genome comparisons are expected to contain some false-positive signals, we combined this approach with traditional QTL mapping, which is more robust to this problem. In order to rapidly screen for genomic regions associated with pelvic phenotype, we began by using bulked segregant genotyping (Postlethwait et al. 1994; Cresko et al. 2004) with 192 previously

described microsatellite markers located throughout the genome (Shapiro et al. 2009) (see Methods and Table 1.4). We found that 32 markers showed differential allele frequencies between complete and reduced pools in at least two families. These candidate markers were used to genotype all individual fish from all four families.

QTL mapping. F1 offspring from all families were genotyped individually with the 32 candidate markers identified by bulked segregant analysis. QTL analysis of genotypes and pelvic phenotypes with GridQTL (Seaton 2006) identified 14 linkage groups that affect at least one component of the pelvic skeleton (Table 1.6). Seven linkage groups (1a, 1b, 3, 8, 12, 14a, and 17) were identified in more than one family, while another seven linkage groups (2, 4, 10, 15b, 16, 18, and 19) affected only one component of the pelvis in a single family. We also note that some linkage groups affect one pelvic structure in one family, and another structure in a second family. For example, linkage group 1b affects left spine and left girdle length in cross 2010-05 but ascending process height in cross 2010-03. These differences among crosses could be due to low numbers of offspring in some families, which would make small-effect QTL difficult to identify, an effect of the differences in the distribution of pelvic phenotype across families or differences in the genetic backgrounds of the parents, which differ between families.

Overall, in the Salt River population, we detected genomic regions of a smaller effect than the major loci previously observed in other crosses (between 4.3 and 25.1 percent variance explained [PVE]). This is notable because, to date, examination of the genetic architecture of pelvic reduction across populations of both threespine and ninespine sticklebacks have found variation in pelvic phenotype to be controlled

primarily by a single, large-effect, QTL (PVE: 59.0-87.0 %) and between 1 and 4 secondary QTL (PVE: 5.6-33.2%) (Shapiro et al. 2004; Coyle et al. 2007; Shapiro et al. 2009; Shikano et al. 2013). Furthermore, although some of the genomic regions implicated in our QTL mapping have been seen in other stickleback populations (linkage groups 4, 7b, and 8), these genomic regions may be acting on pelvic phenotype in a way that is distinct from in other populations. For example, while linkage groups 4 and 8 have been identified as large-effect QTLs in ninespine and threespine stickleback populations, respectively, they have only a small effect on pelvic phenotype in Salt River (LG4, 11.9 PVE; LG8, 4.9-5.9 PVE). Finally, we have also identified a number of novel small-effect QTL that contribute to pelvic reduction in the Salt River population including linkage groups 1b, 3, 10, 12, 14a, 15b, 16, 17, 18, and 19.

Overlap of Pooled Resequencing and QTL Mapping

Of the 14 genomic regions we identified that affect pelvic phenotype in Salt River, 9 were located within 500 kb of a window in the top 0.1% of LRT values. In order to identify genes that might affect pelvic phenotype in Salt River, we compiled a list of candidate genes that were located within 50 kb up- or downstream of a window in the top 0.1% of LRT values that was also within 500 kb of a microsatellite identified in QTL mapping. This list included a total of 12 genes on 3 linkage groups (Table 1.7). As there has only been one gene previously implicated in stickleback pelvic reduction, it is unsurprising that the candidates identified by our study are novel. Furthermore, while some were located in genomic regions previously implicated in pelvic phenotype (linkage group 8, spine length in threespine sticklebacks) (Peichel et al. 2001), others were not (those on linkage groups 3 and 12). Of the list of 12 candidates, two stand out because of

their role in sonic hedgehog (*Shh*) signaling (EFCAB7) and skeletal development (Chst11). EFCAB7 is interesting as its depletion has recently been shown to impair *Shh* signaling in skeletal tissues and mimic Ellis van Creveld syndrome, which is characterized by, among other phenotypes, shortened limbs (Pusapati et al. 2014). Chst11 homozygous null mice exhibit neonatal lethality, dwarfism, and abnormal skeletal structures. Taken together, the combination of traditional QTL mapping and whole-genome resequencing allowed us to characterize the genetic basis of pelvic reduction in Salt River and Pine Lake as a combination of multiple loci of relatively small to moderate effect and also identify at least two noteworthy candidate genes that can easily be assessed in terms of coding or, in the future, expression differences between the two phenotypes.

Linkage Group 12 and Pelvic Phenotype

Using both whole-genome resequencing and QTL mapping techniques, we identified linkage group 12 as a contributor to pelvic phenotype. In an Alaskan population of ninespine sticklebacks lateral plate number maps to this linkage group (Shapiro et al. 2009). Additionally, this linkage group contains the sex-determination region of the genome in ninespine sticklebacks (Ross et al. 2009; Shapiro et al. 2009) and when we compared the overall pelvic score between males and females in Salt River crosses, we found that females have significantly lower pelvic scores than males (ANOVA; p-value < 0.001) (Figures 1.5 and 1.6). This difference between the sexes was seen in all crosses and in all pelvic structures with the exception of the left ascending process.

Because we found that pelvic phenotype differed between the sexes, we tested for

potential differences between QTL identified in males and females using only the largest cross (2008-01). In doing so, we found that the QTL on linkage group 12 is only present in females. This could be because males have an XY genotype at the sex-determining region and would therefore show no heterozygosity at markers in this region; without heterozygosity at a marker, it would not be possible to identify a QTL. We also found three QTL that were specific to males (linkage groups 1a, 5, and 18). This may be because, in general, females are more likely to be missing components of the pelvis. Therefore, QTL found in exclusively female samples could be explaining presence or absence of a structure, while those identified in exclusively male populations account for variation in a structure that is present.

Motivated by these results, we tested for correlations between sex and pelvic phenotype in wild-caught fish. For both the Pine Lake and Salt River populations, we used the first 100 pelvic-complete fish that we collected for resequencing. Likewise, we used the first 62–86 pelvic reduced fish that we collected. Therefore, we do not have a random sample of the population as a whole, but fish were taken randomly from the population within each pelvic phenotype. We found that within wild-caught Salt River fish, there are significantly more females with a reduced pelvis and significantly more males with a complete pelvis (Fisher's exact test, p-value <0.01) (Table 1.8). This pattern is also seen in Salt River F1 offspring: the mean pelvic score is significantly lower in females than males (ANOVA, $p<0.001$) Surprisingly, wild-caught fish from Pine Lake do not show any sex-specific differences in pelvic score. Nucleotide diversity in these populations also suggests that there in no significant difference pelvic phenotype between the sexes in Pine Lake (Figure 1.7). Overall linkage group 12 shows increased values of

pi compared to other genomic regions. This is probably because the reference genome was assembled using a female, which would have two X chromosomes; any reads from Y chromosomes in the pooled sequencing data would likely map on top of the X chromosome. This would result in a level of nucleotide diversity on this linkage group. Interestingly, there are differences in pi on LG12 between the two phenotypic pools in Salt River. Pi is higher on LG 12 in the complete pool, which is composed primarily of males. Pi is lower in the reduced (female-dominant) pool. This difference in pi is not seen in Pine Lake, which implies that the ratio of sexes in each phenotypic pool is similar. These results are unexpected given how close, phylogenetically and geographically, these two populations are. However, there is the possibility that while linkage group 12 contributes to pelvic phenotype in both of these populations (which is observed in resequencing data) the specific region of the linkage group is different between the two. We did not conduct QTL mapping in Pine Lake, but it may be possible that a recombination event since the separation of Salt River and Pine Lake populations has separated the regions of LG12 that contribute to pelvic phenotype and sex determination in one of these populations. That is, a hypothetical "pelvic reduction" gene may originally have been linked to the sex-determining region of LG12 but after a recombination event since the split of the two populations moved to the pseudoautosomal region of the chromosome in Pine Lake. One way to determine if pelvic reduction is mapping to distinct regions of linkage group 12 in Salt River and Pine Lake would be a comparison of QTL mapping results between the two populations. Currently, QTL mapping data for Pine Lake is not available. Furthermore, resequencing data are uninformative as recombination is highly reduced in sex chromosomes, thereby elevating

F_{ST} and LRT values across the entire linkage group.

What might drive this difference in pelvic phenotypes between the sexes? A reduced pelvic skeleton could provide a female-specific advantage, for example, an increase in clutch or egg size. Life history traits such as these have long been thought to be a prime target of selection (Mousseau and Roff 1987) but while many studies have surveyed life history traits in threespine sticklebacks (reviewed in Baker 1994), very few have specifically examined differences between individuals or populations that vary in the extent of bony armor. Data comparing the clutch size of populations with varying lateral plate morphology have been conflicting; while Kynard (1972) reported that females with fewer lateral plates had significantly larger clutches, Baker (1994) reanalyzed the same data and found no correlation. Conversely, Baker et al. (1998) compared female life history traits among 12 Alaskan populations of threespine sticklebacks that varied in pelvic phenotype and found that while clutch size was not statistically different between morphs, mean egg mass was greater in pelvic-reduced populations. Egg size has been shown to be correlated with larger embryos as well as juveniles (Blaxter and Hempel 1963; Reagan and Conley 1977; Thorpe et al. 1984; McKay et al. 1985), an increased growth rate in hatchlings (Wallace and Aasjord 1984), and increased juvenile survival (Marsh 1986). Therefore, egg size is an important life history trait that could impact an individual's lifetime fecundity and fitness.

Differentiation Between Salt River and Pine Lake Populations

Salt River and Pine Lake are located within 40 km of one another and may have been connected at some point in the recent past (Nelson and Paetz 1974), yet these bodies of water differ in several fundamental respects. In addition to typical differences between

lake and stream habitats (e.g., depth, vegetation, water movement), there are also large differences in salinity (Salt River = 2%; Pine Lake = 0.31%) (Nelson 1972). Furthermore, ninespine stickleback residents of each of these habitats differ in overall body morphology and may occupy distinct niches. Fish from Salt River have a more "benthic" appearance, that is, shorter and deeper bodies with shorter fins (personal observation). In contrast, those in Pine Lake appear more "limnetic" with long, streamlined bodies and heads, longer fins, and a narrower caudal region. This combination of traits suggests an open-water niche (Webb 1982; Walker 1997; Walker and Bell 2000; Spoljaric and Reimchen 2007). Because of these morphological and potential physiological dissimilarities, we also examined genomic differentiation between the two populations as a whole to identify genes that may potentially be under selection in these very different habitats.

Both F_{ST} and LRT showed broadly similar patterns and, as expected, overall values of both of these metrics were higher than in within-population comparisons (mean F_{ST} and LRT were 0.053 and 3.55, respectively) (Figure 1.8). Similar genomic comparisons have been done in threespine sticklebacks and have identified genomic regions that differ significantly between marine and freshwater populations as well as benthic and limnetic morphs within a lake (Hohenlohe et al. 2010; Jones et al. 2012a; Jones et al. 2012b). While it is possible that our work may not be directly comparable to previous studies using a different species, any regions that overlap between the two may provide information about recurrent selection at similar genomic regions between species. Collectively, previous work has identified hundreds of SNPs with differing allele frequencies between habitats and nearly every linkage group contains regions of

increased differentiation. However, a handful of linkage groups have been identified repeatedly in multiple studies using SNP genotyping arrays, RAD Tag genotyping, and whole-genome resequencing (1, 4, 7, 11, 20, and 21). The genomic regions with the highest differentiation identified in our study are located on linkage groups 5, 6, 8, 10, 13, 14, 16, and 20. Hohenlohe et al. (2010) identified regions of linkage groups 2, 4, 9, 11, 16, and 19 that were highly differentiated among freshwater populations of threespine sticklebacks, but with the exception of linkage group 16, there is very little overlap between this interpopulation comparison and our own. This could be due to the fact that all populations included in the threespine stickleback comparison were from lake populations and may be detecting selection on genomic regions that would be beneficial in those specific habitats. Because previous work has been conducted in a different species and focused primarily on differences between marine and freshwater habitats (both populations included in our comparison are from freshwater habitats), the lack of overlap between the two are unsurprising. The best comparison to our study may be between benthic and limnetic threespine stickleback species pairs as, morphologically, Salt River resembles a benthic form while Pine Lake appears more limnetic. Jones et al. (2012a) identified 15 genomic regions that differed between benthics and limnetics in multiple lakes, including portions of linkage groups 1, 2, 4, 7, 10, 11, 12, 20, and 21. Again, we found only one linkage group (10) in common between stream and lake populations of ninespine sticklebacks in our study and the comparison between benthic and limnetic threespine stickleback species pairs (Jones et al. 2012a).

The overall lack of overlap between our study and other published work could be that in all other comparisons the authors were specifically looking for regions that were

under selection across multiple populations. Because our work focused on just two populations, we may be seeing signals of differentiation that are very population-specific and would not be picked up in studies such as those done previously. It could also be that the genomic regions under selection in ninespine sticklebacks after a shift to freshwater are different than those in threespine sticklebacks. It has already been noted that different genomic regions control sex, lateral plates, and pelvic reduction in the two species (Peichel et al. 2001; Colosimo et al. 2004; Shapiro et al. 2004; Shapiro et al. 2009). Finally, although both threespine and ninespine sticklebacks have adapted to superficially similar habitats and undergone a similar set of morphological changes associated with a shift to freshwater because they are distinct species with unique natural histories, they simply may not be directly comparable.

Candidate genes in regions of differentiation. In other between-population comparisons in threespine sticklebacks, specific genes have been identified that show repeated selection between benthic and limnetic species pairs (IGK, KITLG, THUMPD3) and marine and freshwater populations (WNT7B, ATPase, EDA, Mucin, SULT4A) (Jones et al. 2012a; Jones et al. 2012b). To test whether any of the same genes might show signatures of selection between Salt River and Pine Lake, we examined genes that were near genomic regions that showed the highest differentiation between the two populations. We compiled a list of genes that were found within 50 kb up- or downstream of peaks that had an LRT score over 60 (top 0.5% of scores). A total of 79 windows across 10 scaffolds met this criterion and contained a total of 45 genes. Interestingly, many genes on this list play a role in immune function (JAK2, CD274, AC3H2, SIGLEC15, SLAMF7). In threespine sticklebacks, even benthic and limnetic species

pairs in the same lake differ in parasite load and populations show adaptation to local parasites (MacColl 2009; Eizaguirre et al. 2012); therefore, it is expected that the Salt River and Pine Lake populations would adapt to distinct local immunological challenges. Other genomic scans in threespine sticklebacks that compared marine and freshwater populations and benthic and limnetic species pairs also identified genes associated with immune function (Jones et al. 2012a; Jones et al. 2012b). Another notable gene that was near a peak of differentiation is DKK1. This gene has been shown to play a role in face and head morphogenesis (Roessler et al. 2000; Mao et al. 2001; Mukhopadhyay et al. 2001) and is an interesting candidate given that a primary morphological difference between Salt River and Pine Lake fish is craniofacial shape (personal observation). In conclusion, while we did not identify any of the same specific genes that show high signatures of selection between benthic and limnetic threespine stickleback populations, we did find some genes that are in the same class as those identified threespine sticklebacks (i.e., immune function). We also identified genes such as DKK1, which have not been noted in other species but are interesting because of their potential role in a morphological difference that characterizes Salt River and Pine Lake.

Conclusions

Because of their dramatic morphological, physiological, and behavioral variation between populations, ninespine sticklebacks provide an ideal model to examine convergent, and adaptively relevant, skeletal phenotypes in wild populations. By using this species to better understand the genetic architecture underlying a dramatic change such as pelvic reduction, we can gain a better understanding of the patterns that underlie evolutionary change in general. For example, how many genetic changes are responsible

for large morphological changes, and do the same genetic changes underlie the repeated evolution of similar traits in different lineages? The work presented here suggests that pelvic reduction is not always explained by a small number of genes of large effect and that of the populations of ninespine sticklebacks that have been examined to date; there are at least three distinct genetic mechanisms that could lead to pelvic reduction.

Additionally, because a considerable amount of work has been done on the genetics of phenotypic variation in the closely related threespine stickleback, work in ninespine sticklebacks adds to the understanding of convergent traits on a micro- and macroevolutionary scale. While there may not be enough populations examined in either species to draw definitive conclusions, this work, combined with current literature, suggests that while there are some mechanisms that underlie pelvic reduction in both species (i.e., *Pitx1*), ninespine sticklebacks have exhibited a broader range of genetic possibilities for pelvic loss.

Here we have presented the draft genome of the ninespine stickleback, which was used as the reference for comparative resequencing of two Canadian populations that display an unusually broad range of pelvic phenotypes. This allowed us to identify a larger than expected number of differentiated genomic regions between pools of individuals with divergent pelvic phenotypes. We were then able to compare patterns of differentiation between these two populations with similar population-level pelvic phenotypes. We found that linkage group 12, the sex-determining linkage group, shows elevated levels of differentiation in both Salt River and Pine Lake. In addition, other regions of the genome show differentiation in both populations. These results suggest that similar genetic mechanisms are responsible for pelvic reduction in these two populations.

To compliment the whole-genome resequencing we also conducted QTL mapping using four half-sibling families from Salt River. These results implicated 14 different genomic regions that contribute to pelvic phenotype. Among these were nine that were located within 500 kb of a window identified as being highly differentiated by whole-genome resequencing. This combination of techniques allowed us to compile a list of candidate genes possibly contributing to pelvic reduction, including candidates that have been implicated in skeletal development and *Shh* signaling.

Overall, the genetic architecture of pelvic reduction that we have described for Salt River and Pine Lake contrasts with previously published studies. Previous work in both threespine and ninespine sticklebacks suggest that pelvic reduction is primarily caused by a few genes of large effect with a small number of modifiers, and in one study a specific molecular change in *Pitx1* was found in multiple populations (Chan et al. 2010). While identification of individual genes controlling phenotype is important and allows for a deeper understanding of the developmental pathways that lead to specific morphologies as well as the selective forces that may be acting on individual alleles in wild populations, simply understanding the genetic architecture of a dramatic and ecologically relevant trait can add to our knowledge of the overall patterns of evolution. Our current knowledge of the genetic basis of large morphological changes may be biased because genes of large effect are easier to identify. Furthermore, once a gene is identified, it is added to a list of candidate genes for future studies on a phenotype and, therefore, may be overrepresented in subsequent literature. Complementation tests or gene expression analyses, for example, are testing for possible effects of previously identified genes or genomic regions (Cole et al. 2003; Shapiro et al. 2009). Other work

has used mapping crosses to examine only a subset of genomic markers in the vicinity of *Pitx1* (Coyle et al. 2007). The work presented here suggests that pelvic reduction in sticklebacks does not always have a genetically simple basis.

Before sticklebacks were used for molecular genetic studies, it had already been noted that pelvic reduction, at least in ninespine sticklebacks, might have different genetic architectures between populations. For example, Ziuganov and Zotin (1995) hypothesize that pelvic reduction in Levin Navolok Creek (Russia) is the result of a single genomic region with complete genetic dominance. In contrast, Blouw and Boyd (1992) found that a polygenic model with a genotypic threshold best explains pelvic reduction in O'Keefe's Lake (Prince Edward Island, Canada). Therefore, while there there are many cases of ecologically relevant traits in sticklebacks and other organisms that have been attributed to a small number of genes or genetic regions, the work presented here is one of a growing number of examples of a complex genetic basis describing a major phenotypic change. See Orr and Coyne (1992) and Rockman (2012) for a review of the possible overrepresentation of large effect QTL in the literature to date.

Furthermore, we identified several QTL that have never been shown to affect pelvic phenotype in other populations. Linkage group 12 is a particularly interesting QTL as this linkage group is the sex-determining region of the genome. This suggests that pelvic phenotype is correlated with sex in this population. In fact, we did see significantly lower pelvic scores in wild-caught females from Salt River as well as females from crosses, but did not see that same pattern in wild-caught Pine Lake fish, suggesting a recent recombination event in the Pine Lake population. Differences in pelvic phenotype

between the sexes has not been previously described and raises the possibility that pelvic reduction could have an affect on female-specific reproductive traits such as clutch size. Possible differences in reproductive traits such as increased clutch or egg size in fish with lower amounts of armor have been reported, but those studies were done comparing armor phenotypes from different populations. More work will need to be done to see if this pattern is the case *within* the Salt River and Pine Lake populations. If pelvic reduction is found to be correlated with differences in reproductively relevant traits, it could mean that in addition to the current hypotheses (including differences in calcium availability and the presence of grasping predators) increased fecundity may also need to be considered a possible selective force driving pelvic reduction.

References

Aldenhoven, J. T., M. A. Miller, P. S. Corneli, and M. D. Shapiro. 2010. Phylogeography of ninespine sticklebacks (Pungitius pungitius) in North America: glacial refugia and the origins of adaptive traits. Mol Ecol 19:4061-4076.

Altschul, S. F., W. Gish, W. Miller, E. W. Myers, and D. J. Lipman. 1990. Basic local alighnment search tool. J Mol Biol 215:403-410.

Baker, J. A. 1994. Life history variation in female threespine stickleback. Pp. 144-187 *in* M. A. Bell, and S. A. Foster, eds. The Evolutionary Biology of the Threespine Stickleback. Oxford University Press, New York.

Baker, J. A., S. A. Foster, D. C. Heins, M. A. Bell, and R. W. King. 1998. Variation in female life-history traits among Alaskan populations of the threespine stickleback, Gasterosteus aculeatus L. (Pisces: Gasterosteidae). Biol J Linn Soc Lond 63:141-159.

Bell, A. M., G. Orti, J. A. Walker, and J. P. Koenings. 1993. Evolution of pelvic reduction in threespine stickleback fish-a test of competing hypotheses. Evolution 47:906-914.

Bell, M. A. 1987. Interacting evolutionary constrains in pelvic reduction of threespine sticklebacks, *Gasterosteus aculeatus* (Pisces, Gasterosteidae). Biol J Linn Soc 31:347-382.

Bell, M. A. and S. A. Foster. 1994. The Evolutionary Biology of the Threespine Stickleback. Oxford Univ Press, Oxford.

Bell, M. A. and G. Orti. 1994. Pelvic reduction in threespine stickleback from Cook Inlet lakes: geographic distribution and intrapopulation variation. Copeia 1994:314-325.

Bernatchez, L. and C. C. Wilson. 1998. Comparative phylogeography of Nearctic and Palearctic fishes. Molecular Ecology 7:431-452.

Blaxter, J. H. S. and G. Hempel. 1963. The influence of egg size on herring larvae, *Clupea harengus*. J. Cons. Inter. Explor. Mer. 28:211-240.

Blouw, D. M. and G. J. Boyd. 1992. Inheritance of reduction, loss, and asymmetry of the pelvis of *Pungitius pungitius* (ninespine stickleback). Heredity 68:33-42.

Burke, M. K., J. P. Dunham, P. Shahrestani, K. R. Thornton, M. R. Rose, and A. D. Long. 2010. Genome-wide analysis of a long-term evolution experiment with Drosophila. Nature 467:587-590.

Burridge, C. P., D. Craw, D. Fletcher, and J. M. Waters. 2008. Geological dates and molecular rates: fish DNA sheds light on time dependency. Mol Biol Evol 25:624-633.

Campbell, M. S., M. Law, C. Holt, J. C. Stein, G. D. Moghe, D. E. Hufnagel, J. Lei, R. Achawanantakun, D. Jiao, C. J. Lawrence, D. Ware, S. H. Shiu, K. L. Childs, Y. Sun, N. Jiang, and M. Yandell. 2014. MAKER-P: A Tool Kit for the Rapid Creation, Management, and Quality Control of Plant Genome Annotations. Plant physiology 164:513-524.

Cantarel, B. L., I. Korf, S. M. Robb, G. Parra, E. Ross, B. Moore, C. Holt, A. Sanchez Alvarado, and M. Yandell. 2008. MAKER: an easy-to-use annotation pipeline designed for emerging model organism genomes. Genome Res 18:188-196.

Chan, Y. F., M. E. Marks, F. C. Jones, G. Villarreal, Jr., M. D. Shapiro, S. D. Brady, A. M. Southwick, D. M. Absher, J. Grimwood, J. Schmutz, R. M. Myers, D. Petrov, B. Jonsson, D. Schluter, M. A. Bell, and D. M. Kingsley. 2010. Adaptive evolution of pelvic reduction in sticklebacks by recurrent deletion of a Pitx1 enhancer. Science 327:302-305.

Cole, N. J., M. Tanaka, A. Prescott, and C. Tickle. 2003. Expression of limb initiation genes and clues to the morphological diversification of threespine stickleback. Curr Biol 13:R951-952.

Colosimo, P. F., C. L. Peichel, K. Nereng, B. K. Blackman, M. D. Shapiro, D. Schluter, and D. M. Kingsley. 2004. The genetic architecture of parallel armor plate reduction in threespine sticklebacks. PLoS Biol 2:E109.

Coyle, S. M., F. A. Huntingford, and C. L. Peichel. 2007. Parallel evolution of Pitx1 underlies pelvic reduction in Scottish threespine stickleback (Gasterosteus aculeatus). J Hered 98:581-586.

Cresko, W. A., A. Amores, C. Wilson, J. Murphy, M. Currey, P. Phillips, M. A. Bell, C. B. Kimmel, and J. H. Postlethwait. 2004. Parallel genetic basis for repeated evolution of armor loss in Alaskan threespine stickleback populations. Proc Natl Acad Sci U S A 101:6050-6055.

Earl, D. A. and B. M. vonHoldt. 2011. STRUCTURE HARVESTER: a website and program for visualizing STRUCTURE output and implementing the Evanno method. Conserv. Genet. Resour. 4:359–361.

Eizaguirre, C., T. L. Lenz, M. Kalbe, and M. Milinski. 2012. Divergent selection on locally adapted major histocompatibility complex immune genes experimentally proven in the field. Ecol Lett 15:723-731.

Evanno, G., S. Regnaut, and J. Goudet. 2005. Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. Mol Ecol 14:2611-2620.

Grabherr, M. G., B. J. Haas, M. Yassour, J. Z. Levin, D. A. Thompson, I. Amit, X. Adiconis, L. Fan, R. Raychowdhury, Q. Zeng, Z. Chen, E. Mauceli, N. Hacohen, A. Gnirke, N. Rhind, F. di Palma, B. W. Birren, C. Nusbaum, K. Lindblad-Toh, N. Friedman, and A. Regev. 2011. Full-length transcriptome assembly from RNA-Seq data without a reference genome. Nature biotechnology 29:644-652.

Gross, H. P. 1978. Natural selection by predators on the defensive apparatus of the three-spined stickleback, *Gasterosteus aculeatus* L. Can J Zool 56:398-413.

Hagen, D. W. and L. G. Gilbertson. 1972. Geographic variation and environmental selection in *Gasterosteus aculeaus* L. in the Pacific northwest, America. Evolution 26:32-51.

Herczeg, G., M. Turtiainen, and J. Merilä. 2010. Morphological divergence of North-European nine-spined sticklebacks (*Pungitius pungitius*): signatures of parallel evolution. Biol J Linn Soc 101:403-416.

Hewitt, G. 2000. The genetic legacy of the Quaternary ice ages. Nature 405:907-913.

Hohenlohe, P. A., S. Bassham, P. D. Etter, N. Stiffler, E. A. Johnson, and W. A. Cresko. 2010. Population genomics of parallel adaptation in threespine stickleback using sequenced RAD tags. PLoS Genet 6:e1000862.

Holt, C. and M. Yandell. 2011. MAKER2: an annotation pipeline and genome-database management tool for second-generation genome projects. BMC Bioinformatics

12:491.

Holt, R. A., G. M. Subramanian, A. Halpern, G. G. Sutton, R. Charlab, D. R. Nusskern, P. Wincker, A. G. Clark, J. M. Ribeiro, R. Wides, S. L. Salzberg, B. Loftus, M. Yandell, W. H. Majoros, D. B. Rusch, Z. Lai, C. L. Kraft, J. F. Abril, V. Anthouard, P. Arensburger, P. W. Atkinson, H. Baden, V. de Berardinis, D. Baldwin, V. Benes, J. Biedler, C. Blass, R. Bolanos, D. Boscus, M. Barnstead, S. Cai, A. Center, K. Chaturverdi, G. K. Christophides, M. A. Chrystal, M. Clamp, A. Cravchik, V. Curwen, A. Dana, A. Delcher, I. Dew, C. A. Evans, M. Flanigan, A. Grundschober-Freimoser, L. Friedli, Z. Gu, P. Guan, R. Guigo, M. E. Hillenmeyer, S. L. Hladun, J. R. Hogan, Y. S. Hong, J. Hoover, O. Jaillon, Z. Ke, C. Kodira, E. Kokoza, A. Koutsos, I. Letunic, A. Levitsky, Y. Liang, J. J. Lin, N. F. Lobo, J. R. Lopez, J. A. Malek, T. C. McIntosh, S. Meister, J. Miller, C. Mobarry, E. Mongin, S. D. Murphy, D. A. O'Brochta, C. Pfannkoch, R. Qi, M. A. Regier, K. Remington, H. Shao, M. V. Sharakhova, C. D. Sitter, J. Shetty, T. J. Smith, R. Strong, J. Sun, D. Thomasova, L. Q. Ton, P. Topalis, Z. Tu, M. F. Unger, B. Walenz, A. Wang, J. Wang, M. Wang, X. Wang, K. J. Woodford, J. R. Wortman, M. Wu, A. Yao, E. M. Zdobnov, H. Zhang, Q. Zhao, S. Zhao, S. C. Zhu, I. Zhimulev, M. Coluzzi, A. della Torre, C. W. Roth, C. Louis, F. Kalush, R. J. Mural, E. W. Myers, M. D. Adams, H. O. Smith, S. Broder, M. J. Gardner, C. M. Fraser, E. Birney, P. Bork, P. T. Brey, J. C. Venter, J. Weissenbach, F. C. Kafatos, F. H. Collins and S. L. Hoffman. 2002. The genome sequence of the malaria mosquito Anopheles gambiae. Science 298:129-149.

Hoogland, R. D., D. Morris, and N. Tinbergen. 1957. The spines of sticklebacks (*Gasterosteus* and *Pygosteus*) as means of defense against predators (*Perca* and *Esox*). Behaviour 10:205-230.

Jaillon, O., J. M. Aury, F. Brunet, J. L. Petit, N. Stange-Thomann, E. Mauceli, L. Bouneau, C. Fischer, C. Ozouf-Costaz, A. Bernot, S. Nicaud, D. Jaffe, S. Fisher, G. Lutfalla, C. Dossat, B. Segurens, C. Dasilva, M. Salanoubat, M. Levy, N. Boudet, S. Castellano, V. Anthouard, C. Jubin, V. Castelli, M. Katinka, B. Vacherie, C. Biemont, Z. Skalli, L. Cattolico, J. Poulain, V. De Berardinis, C. Cruaud, S. Duprat, P. Brottier, J. P. Coutanceau, J. Gouzy, G. Parra, G. Lardier, C. Chapple, K. J. McKernan, P. McEwan, S. Bosak, M. Kellis, J. N. Volff, R. Guigo, M. C. Zody, J. Mesirov, K. Lindblad-Toh, B. Birren, C. Nusbaum, D. Kahn, M. Robinson-Rechavi, V. Laudet, V. Schachter, F. Quetier, W. Saurin, C. Scarpelli, P. Wincker, E. S. Lander, J. Weissenbach, and H. Roest Crollius. 2004. Genome duplication in the teleost fish Tetraodon nigroviridis reveals the early vertebrate proto-karyotype. Nature 431:946-957.

Janssen, S., G. Ramaswami, E. E. Davis, T. Hurd, R. Airik, J. M. Kasanuki, L. Van Der Kraak, S. J. Allen, P. L. Beales, N. Katsanis, E. A. Otto, and F. Hildebrandt. 2011. Mutation analysis in Bardet-Biedl syndrome by DNA pooling and massively parallel resequencing in 105 individuals. Hum Genet 129:79-90.

Jones, F. C., Y. F. Chan, J. Schmutz, J. Grimwood, S. D. Brady, A. M. Southwick, D. M. Absher, R. M. Myers, T. E. Reimchen, B. E. Deagle, D. Schluter, and D. M. Kingsley. 2012a. A genome-wide SNP genotyping array reveals patterns of global and repeated species-pair divergence in sticklebacks. Curr Biol 22:83-90.

Jones, F. C., M. G. Grabherr, Y. F. Chan, P. Russell, E. Mauceli, J. Johnson, R. Swofford, M. Pirun, M. C. Zody, S. White, E. Birney, S. Searle, J. Schmutz, J. Grimwood, M. C. Dickson, R. M. Myers, C. T. Miller, B. R. Summers, A. K. Knecht, S. D. Brady, H. Zhang, A. A. Pollen, T. Howes, C. Amemiya, J. Baldwin, T. Bloom, D. B. Jaffe, R. Nicol, J. Wilkinson, E. S. Lander, F. Di Palma, K. Lindblad-Toh, and D. M. Kingsley. 2012b. The genomic basis of adaptive evolution in threespine sticklebacks. Nature 484:55-61.

Kim, S. Y., Y. Li, Y. Guo, R. Li, J. Holmkvist, T. Hansen, O. Pedersen, J. Wang, and R. Nielsen. 2010. Design of association studies with pooled or un-pooled next-generation sequencing data. Genet Epidemiol 34:479-491.

Klepaker, T., K. Ostbye, and M. A. Bell. 2013. Regressive evolution of the pelvic complex in stickleback fishes: a study of convergent evolution. Evol Ecol Res 15:1-23.

Korf, I. 2004. Gene Finding in Novel Genomes. BMC Bioinformatics 5:59–67.

Korf, I., M. Yandell, and J. Bedel. 2003. BLAST. O'Reily, Cambridge.

Kynard, B. E. 1972. Male breeding behavior and lateral plate phenotypes in the threespine stickleback (Gasterosteus aculeatus L.). University of Washington, Seattle.

Langmead, B. and S. L. Salzberg. 2012. Fast gapped-read alignment with Bowtie 2. Nature methods 9:357-359.

Lescak, E. A. and F. A. von Hippel. 2011. Selective predation of threespine stickleback by rainbow trout. Ecology of Freshwater Fish 20:308-314.

MacColl, A. D. C. 2009. Parasite burdens differ between sympatric three-spined stickleback species. Ecography 32:153-160.

Mao, B., W. Wu, Y. Li, D. Hoppe, P. Stannek, A. Glinka, and C. Niehrs. 2001. LDL-receptor-related protein 6 is a receptor for Dickkopf proteins. Nature 411:321-325.

Marcais, G. and C. Kingsford. 2011. A fast, lock-free approach for efficient parallel counting of occurrences of k-mers. Bioinformatics 27:764-770.

Marchinko, K. B. 2009. Predation's role in repeated phenotypic and genetic divergence of

armor in threespine stickleback. Evolution 63:127-138.

Marklund, S. and O. Carlborg. 2010. SNP detection and prediction of variability between chicken lines using genome resequencing of DNA pools. BMC Genomics 11:665.

Marsh, E. 1986. Effects of Egg Size on Offspring Fitness an Maternal Fecundity in the Orangethroat Darter, *Etheostoma spectabile* (Pisces: Percidae). Copeia 1986:18-30.

McKay, L. R., P. E. Ihssen, and G. W. Friars. 1985. Genetic parameters of growth in rainbow trout, *Salmo gairdneri*, prior to maturation. Can. J. Genet. Cyto 28:306-312.

Mobley, K. B., D. Lussetti, F. Johansson, G. Englund, and F. Bokma. 2011. Morphological and genetic divergence in Swedish postglacial stickleback (Pungitius pungitius) populations. BMC Evol Biol 11:287.

Moodie, G. E. E. 1972. Predation, natural selection and adaptation in an unusual threespine stickleback. Heredity 28:155-167.

Mousseau, T. A. and D. A. Roff. 1987. Natural selection and the heritability of fitness components. Heredity (Edinb) 59 (Pt 2):181-197.

Mukhopadhyay, M., S. Shtrom, C. Rodriguez-Esteban, L. Chen, T. Tsukui, L. Gomer, D. W. Dorward, A. Glinka, A. Grinberg, S.-P. Huang, C. Niehrs, J. C. I. Belmonte, and H. Westphal. 2001. Dickkopf1 is required for embryonic head induction and limb morphogenesis in the mouse. Dev. Cell 1:423-434.

Near, T. J., R. I. Eytan, A. Dornburg, K. L. Kuhn, J. A. Moore, M. P. Davis, P. C. Wainwright, M. Friedman, and W. L. Smith. 2012. Resolution of ray-finned fish phylogeny and timing of diversification. Proc Natl Acad Sci U S A 109:13698-13703.

Nelson, J. S. 1971. Absence of the pelvic complex in ninespine sticklebacks, *Pungitius pungitius*, collected in Ireland and Wood Buffalo National Park region, Canada, with notes on meristic variation. Copeia:707-717.

Nelson, J. S. and F. M. Atton. 1971. Georgraphic and morphological variation in the presence and absence of the pelvic skeleton in the brook stickleback, Culaea inconstans (Kirtland), in Alberta and Saskatchewan. Can J Zool 49:343-352.

Nelson, J. S. and M. J. Paetz. 1972. Fishes of the north-eastern Wood Buffalo National Park region, Alberta and North-West Territories. Can Field-Nat 86:133-144.

Orr, H. A. and J. A. Coyne. 1992. The genetics of adaptation: a reassessment. Am Nat 140:725-742.

Parra, G., K. Bradnam, and I. Korf. 2007. CEGMA: a pipeline to accurately annotate core genes in eukaryotic genomes. Bioinformatics 23:1061-1067.

Peichel, C. L., K. S. Nereng, K. A. Ohgi, B. L. Cole, P. F. Colosimo, C. A. Buerkle, D. Schluter, and D. M. Kingsley. 2001. The genetic architecture of divergence between threespine stickleback species. Nature 414:901-905.

Posada, D. and K. A. Crandall. 1998. MODELTEST: testing the model of DNA substitution. Bioinformatics 14:817-818.

Postlethwait, J. H., S. L. Johnson, C. N. Midson, W. S. Talbot, M. Gates, E. W. Ballinger, D. Africa, R. Andrews, T. Carl, J. S. Eisen, and et al. 1994. A genetic linkage map for the zebrafish. Science 264:699-703.

Pritchard, J. K., M. Stephens, and P. Donnelly. 2000. Inference of population structure using multilocus genotype data. Genetics 155:945-959.

Pusapati, G. V., C. E. Hughes, K. V. Dorn, D. Zhang, P. Sugianto, L. Aravind, and R. Rohatgi. 2014. EFCAB7 and IQCE Regulate Hedgehog Signaling by Tethering the EVC-EVC2 Complex to the Base of Primary Cilia. Developmental cell 28:483-496.

Quevillon, E., V. Silventoinen, S. Pillai, N. Harte, N. Mulder, R. Apweiler, and R. Lopez. 2005. InterProScan: protein domains identifier. Nucleic Acids Res 33:W116-120.

Reagan, R. E. and C. M. Conley. 1977. Effect of egg diameter on growth of channel catfish. Prog. Fish Cult. 39:133-134.

Reimchen, T. E. 1980. Spine deficiency and polymorphism in a population of *Gasterosteus aculeatus*-an adaptation to predators. Can J Zool 58:1232-1244.

Reist, J. D. 1980. Predation upon pelvic phenotypes of brook stickleback, Culaea inconstans, by selected invertebrates. Can J Zool 58:1253-1258.

Rellstab, C., S. Zoller, A. Tedder, F. Gugerli, and M. C. Fischer. 2013. Validation of SNP allele frequencies determined by pooled next-generation sequencing in natural populations of a non-model plant species. PLoS One 8:e80422.

Rockman, M. V. 2012. The QTN program and the alleles that matter for evolution: all that's gold does not glitter. Evolution 66:1-17.

Roessler, E., Y. Du, A. Glinka, A. Dutra, C. Niehrs, and M. Muenke. 2000. The genomic structure, chromosome location, and analysis of the human DKK1 head inducer gene as a candidate for holoprosencephaly. Cytogenet Cell Genet 89:220-224.

Ross, J. A., J. R. Urton, J. Boland, M. D. Shapiro, and C. L. Peichel. 2009. Turnover of

sex chromosomes in the stickleback fishes (gasterosteidae). PLoS Genet 5:e1000391.

Rubin, C. J., H. J. Megens, A. Martinez Barrio, K. Maqbool, S. Sayyab, D. Schwochow, C. Wang, O. Carlborg, P. Jern, C. B. Jorgensen, A. L. Archibald, M. Fredholm, M. A. Groenen, and L. Andersson. 2012. Strong signatures of selection in the domestic pig genome. Proc Natl Acad Sci USA 109:19529-19536.

Rubin, C. J., M. C. Zody, J. Eriksson, J. R. Meadows, E. Sherwood, M. T. Webster, L. Jiang, M. Ingman, T. Sharpe, S. Ka, F. Hallbook, F. Besnier, O. Carlborg, B. Bed'hom, M. Tixier-Boichard, P. Jensen, P. Siegel, K. Lindblad-Toh, and L. Andersson. 2010. Whole-genome resequencing reveals loci under selection during chicken domestication. Nature 464:587-591.

Sahana, G., D. J. de Koning, B. Guldbrandtsen, P. Sorensen, and M. S. Lund. 2006. The efficiency of mapping of quantitative trait loci using cofactor analysis in half-sib design. Genetics, selection, evolution : GSE 38:167-182.

Santini, F., L. J. Harmon, G. Carnevale, and M. E. Alfaro. 2009. Did genome duplication drive the origin of teleosts? A comparative study of diversification in ray-finned fishes. BMC Evol Biol 9.

Seaton, G. H., J., Grunchec, J-A.; White, I.; Allen, J.; De Koning, D.J.; Wei, W.; Berry, D.; Haley, C., Knott, S. 2006. GridQTL: A grid portal for QTL mapping of compute intensive datasets. Proceedings of the 8th World Congress on Genetics Applied to Livestock Production, Belo Horizonte, Brazil.

Shapiro, M. D., M. A. Bell, and D. M. Kingsley. 2006a. Parallel genetic origins of pelvic reduction in vertebrates. Proc Natl Acad Sci USA 103:13753-13758.

Shapiro, M. D., Z. Kronenberg, C. Li, E. T. Domyan, H. Pan, M. Campbell, H. Tan, C. D. Huff, H. Hu, A. I. Vickrey, S. C. Nielsen, S. A. Stringham, H. Hu, E. Willerslev, M. T. Gilbert, M. Yandell, G. Zhang, and J. Wang. 2013. Genomic diversity and evolution of the head crest in the rock pigeon. Science 339:1063-1067.

Shapiro, M. D., M. E. Marks, C. L. Peichel, B. K. Blackman, K. S. Nereng, B. Jónsson, D. Schluter, and D. M. Kingsley. 2004. Genetic and developmental basis of evolutionary pelvic reduction in threespine sticklebacks. Nature 428:717-723.

Shapiro, M. D., M. E. Marks, C. L. Peichel, B. K. Blackman, K. S. Nereng, B. Jónsson, D. Schluter, and D. M. Kingsley. 2006b. Corrigendum: Genetic and developmental basis of evolutionary pelvic reduction in threespine sticklebacks. Nature 439.

Shapiro, M. D., B. R. Summers, S. Balabhadra, J. T. Aldenhoven, A. L. Miller, C. B. Cunningham, M. A. Bell, and D. M. Kingsley. 2009. The genetic architecture of skeletal convergence and sex determination in ninespine sticklebacks. Curr Biol

19:1140-1145.

Shikano, T., G. Herczeg, and J. Merila. 2011. Molecular sexing and population genetic inference using a sex-linked microsatellite marker in the nine-spined stickleback (Pungitius pungitius). BMC Res Notes 4:119.

Shikano, T., V. N. Laine, G. Herczeg, J. Vilkki, and J. Merila. 2013. Genetic architecture of parallel pelvic reduction in ninespine sticklebacks. G3 3:1833-1842.

Smit, A. F. A. and R. Hubley. 2008.

Smit, A. F. A., R. Hubley, and P. Green. 1996. RepeatMasker Open-3.0.

Spoljaric, M. A. and T. E. Reimchen. 2007. 10,000 years later: evolution of body shape in Haida Gwaii three-spined stickleback. J Fish Biol 70:1484.

Stanke, M., M. Diekhans, R. Baertsch, and D. Haussler. 2008. Using native and syntenically mapped cDNA alignments to improve de novo gene finding. Bioinformatics 24:637-644.

Stanke, M. and S. Waack. 2003. Gene prediction with a hidden Markov model and a new intron submodel. Bioinformatics 19 Suppl 2:ii215-225.

Stinchcombe, J. R. and H. E. Hoekstra. 2008. Combining population genomics and quantitative genetics: finding the genes underlying ecologically important traits. Heredity 100:158-170.

Thorpe, J. E., M. S. Miles, and D. S. Keay. 1984. Developmental rate, fecundity and egg size in Atlantic salmon, *Salmo salar*. Aquaculture. 43:289-305.

Turner, T. L., E. C. Bourne, E. J. Von Wettberg, T. T. Hu, and S. V. Nuzhdin. 2010. Population resequencing reveals local adaptation of Arabidopsis lyrata to serpentine soils. Nat Genet 42:260-263.

Udpa, N., D. Zhou, G. G. Haddad, and V. Bafna. 2011. Tests of selection in pooled case-control data: an empirical study. Front Genet 2:83.

Van der Auwera, G. A., M. O. Carneiro, C. Hartl, R. Poplin, G. del Angel, A. Levy-Moonshine, T. Jordan, K. Shakir, D. Roazen, J. Thibault, E. Banks, K. V. Garimella, D. Altshuler, S. Gabriel, and M. A. DePristo. 2013. From FastQ data to high-confidence variant calls: The Genome Analysis Toolkit best practices pipeline. Current Protocols in Bioinformatics 43:11.10.11-11.10.33.

Vinson, J. P., D. B. Jaffe, K. O'Neill, E. K. Karlsson, N. Stange-Thomann, S. Anderson, J. P. Mesirov, N. Satoh, Y. Satou, C. Nusbaum, B. Birren, J. E. Galagan, and E. S. Lander. 2005. Assembly of polymorphic genomes: algorithms and application to

Ciona savignyi. Genome Res 15:1127-1135.

Walker, J. A. 1997. Ecological morphology of lacustrine three-spine stickleback *Gasterosteus aculeatus* L. (Gasterosteidae) body shape. Biol J Linn Soc 61:3-50.

Walker, J. A. and M. A. Bell. 2000. Net evolutionary trajectories of body shape evolution within a microgeographic radiation of threespine sticklebacks. J Zool Lond 252:293-302.

Wallace, J. C. and D. Aasjord. 1984. An investigation of the consequences of egg size for the culture of Arctic Charr, *Salvelinus alpinus* (L.) Journal of Fish Biology 24:427-435.

Webb, P. W. 1982. Locomotor patters in the evolution of actinopterygian fishes. Am Zool 22:329-342.

Wei, Z., W. Wang, P. Hu, G. J. Lyon, and H. Hakonarson. 2011. SNVer: a statistical tool for variant calling in analysis of pooled or individual next-generation sequencing data. Nucleic Acids Res 39:e132.

Weir, B. S. and C. C. Cockerham. 1984. Estimating F-Statistics for the Analysis of Population Structure. Evolution 38:1358-1370.

Wootton, R. J. 1976. The Biology of the Sticklebacks. Academic, London.

Yang, Z. 2007. PAML 4: phylogenetic analysis by maximum likelihood. Mol Biol Evol 24:1586-1591.

Zhou, D., N. Udpa, M. Gersten, D. W. Visk, A. Bashir, J. Xue, K. A. Frazer, J. W. Posakony, S. Subramaniam, V. Bafna, and G. G. Haddad. 2011. Experimental selection of hypoxia-tolerant *Drosophila melanogaster*. Proc Natl Acad Sci U S A 108:2349-2354.

Zhu, Y., A. O. Bergland, J. Gonzalez, and D. A. Petrov. 2012. Empirical validation of pooled whole genome population re-sequencing in Drosophila melanogaster. PLoS One 7:e41901.

Ziuganov, V. V. and A. A. Zotin. 1995. Pelvic girdle polymorphism and reproductive barriers in the ninespine stickleback Pungitius pungitius (L.) from northwest Russia. Behaviour 132:1095-1105.

Table 1.1. Summary of genomic libraries used for reference sequence

Library Insert Size	Standard Deviation of Insert Size	Number of Reads	Coverage
250 bp	± 30 bp	132,407,570	33.64x
500 bp	± 60 bp	116,127,466	29.5x
2400 bp	±300 bp	304,565,738	77.37x

Table 1.2. Summary of additional teleost genome assemblies

Species	Assembly	Size (bp)	Protein Coding Genes
Astyanax mexicanus	AstMex102	964,248,202	23,042
Danio rerio	Zv.9	1,505,581,940	26,459
Gadus morhua	gadMor1	608,029,870	20,095
Gasterosteus aculeatus	BROADS1	446,627,861	20,787
Oreochromis niloticus	Orenil 1.0	815,725,529	21,437
Oryzias latipes	MEDAKA1	700,386,597	19,699
Pungitius pungitius	PunPun1	441,103,789	22,432
Takifugu rubripes	FUGU 4.0	393,312,790	18,523
Tetraodon nigroviridis	TETRAODON 8.0	342,419,788	19,602

Table 1.3. Summary of Salt River crosses

Family	Year	Male	Male Pelvic Score	Female	Female Pelvic Score	Offspring (n)
2008-01	2008	Male 08-2	8	Female 08-4	7	45
	2008	Male 08-2	8	Female 08-5	8	53
	2008	Male 08-2	8	Female 08-6	8	31
	2008	Male 08-2	8	Female 08-7	8	49
2010-01	2010	Male 10-1	8	Female 10-1	0	21
	2010	Male 10-1	8	Female 10-3	0	35
2010-03	2010	Male 10-3	8	Female 10-11	8	38
	2010	Male 10-3	8	Female 10-13	0	24
	2010	Male 10-3	8	Female 10-14	0	27
2010-05	2010	Male 10-5	8	Female 10-20	8	15
	2010	Male 10-5	8	Female 10-22	0	24
	2010	Male 10-5	8	Female 10-23	8	19

Table 1.4. Summary of samples used in bulked segregant analysis

Cross Name	n included in complete BSA pool	n included in reduced BSA pool
2010-01	14	16
2010-03	21	22
2010-05	16	11

Table 1.5. Genome metrics

Genome size (bp)	441,103,789
Coverage	140.5 x
Number of contigs	8784
Contig N50 length (bp)	122,764
Mean Contig length (bp)	45,262
Number of scaffolds	7850
Scaffold N50 length (bp)	302,754
Mean scaffold length (bp)	56,191
Exonic sequence (bp)	47,709,142
Intronic sequence (bp)	171,695,097
Intergenic sequence (bp)	221,699,550
Number of Genes	22,432
Median gene length (bp)	5,604
Median exon length (bp)	132
Median intron length (bp)	227

Table 1.6. Summary of genomic regions identified by QTL mapping

		Family			
Linkage Group	**Structure**	**2008-01**	**2010-01**	**2010-03**	**2010-05**
1a	left spine	9.4			6.3
1a	right spine	11.7			
1b	left spine				6.9
1b	left ascending process			5.7	
1b	right ascending process			5.4	
1b	left girdle				5.6
1b	right girdle			5.8	
2	left girdle				4.6
3	left ascending process		9.6		
3	left girdle	4.5	6.9		
4	left ascending process				11.9
8	left spine				6.4
8	right spine		5.9		
8	right ascending process		5.9		
8	left girdle				4.9
8	right girdle		5.6		
10	right girdle				4.2
12	left spine		5.7	5.2	
12	right spine				
12	right girdle	5.9		4.4	
12	left ascending process	7.4			
14a	left spine			5.5	
14a	right ascending process		9.6		
14a	right girdle		5.2		
15b	right ascending process				5.2
15b	right girdle				4.1
16	right spine			5.9	
16	left ascending process			7.9	
16	right ascending process			5.8	
17	right spine	4.3			
17	left girdle			9.9	
17	right girdle	5.2			
18	left spine	7.5			
18	right ascending process	5.5			
19	left girdle			6.9	

F-test: light pink, $p < 0.05$; dark pink, $p < 0.01$; numbers indicate percent variance explained by male genotype as determined by ANOVA.

Table 1.7. Candidate genes of pelvic reduction

Linkage Group	Scaffold	Gene	Abbreviation
3	scaffold1112	Neurogenic differentiation factor 6-B	neurod6b
3	scaffold1112	Calcium/calmodulin-dependent 3'C5'-cyclic nucleotide phosphodiesterase 1C	PDE1C
8	scaffold80	EF-hand calcium-binding domain-containing protein 7	EFCAB7
8	scaffold80	Phosphoglucomutase-1	Pgm1
8	scaffold80	Tyrosine-protein kinase transmembrane receptor ROR1	Ror1
8	scaffold140	DNA-dependent protein kinase catalytic subunit	Prkdc
8	scaffold140	Intestinal-type alkaline phosphatase 1	Alpi
12	scaffold14	Tyrosine-protein phosphatase non-receptor type 11	PTPN11
12	scaffold14	Plexin-A1	PLXNA1
12	scaffold14	Carbohydrate sulfotransferase 11	Chst11
12	scaffold14	RING finger protein 223	RNF223
12	scaffold14	Agrin	AGRN

Table 1.8. Pelvic phenotypes by sex in wild-caught fish

	Female	Male
SALT RIVER – complete*	38	60
SALT RIVER – reduced*	42	18
PINE LAKE – complete	43	54
PINE LAKE – reduced	45	37

Fisher's exact test (*P < 0.01)

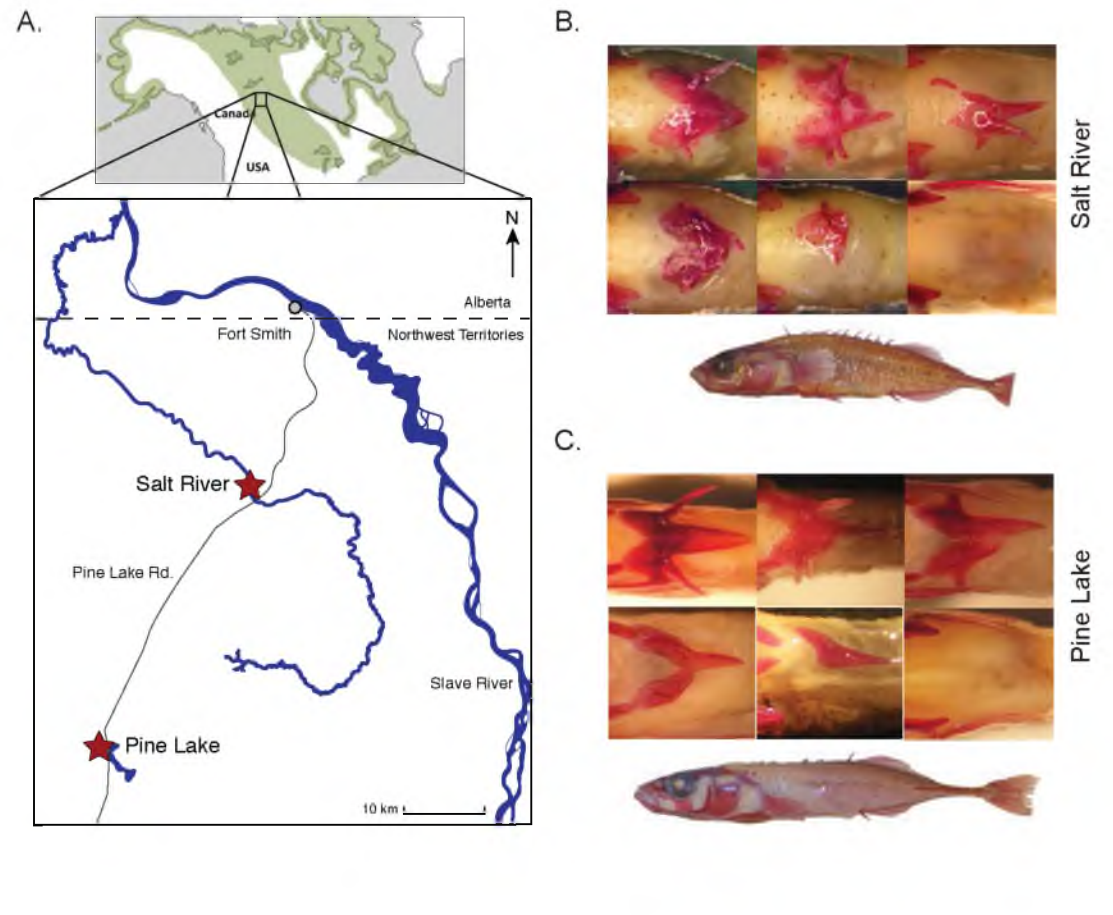


Figure 1.1. Collection sites and variation in phenotype in Salt River and Pine Lake. A) Locations of sites where fish were collected for both whole-genome resequencing and QTL mapping crosses. B) Variation in pelvic phenotype (top) and whole body shape (bottom) seen in both Salt River and C) Pine Lake (right). In both of these populations, there are individuals with pelvic phenotypes ranging from complete (top, left) to absent (bottom, right)

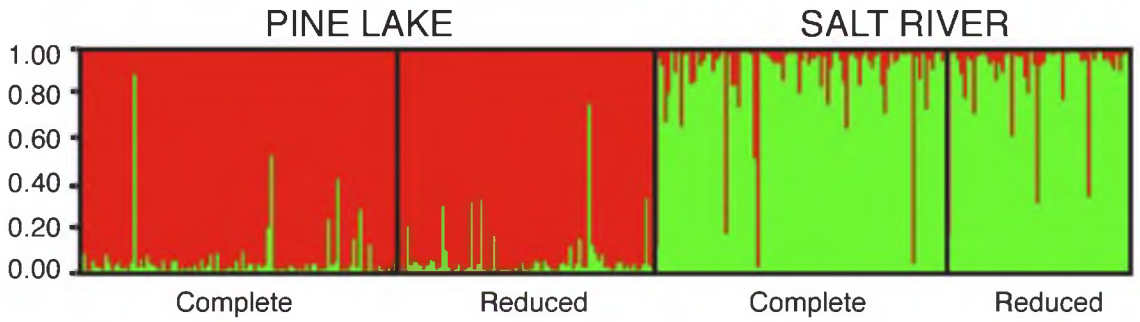
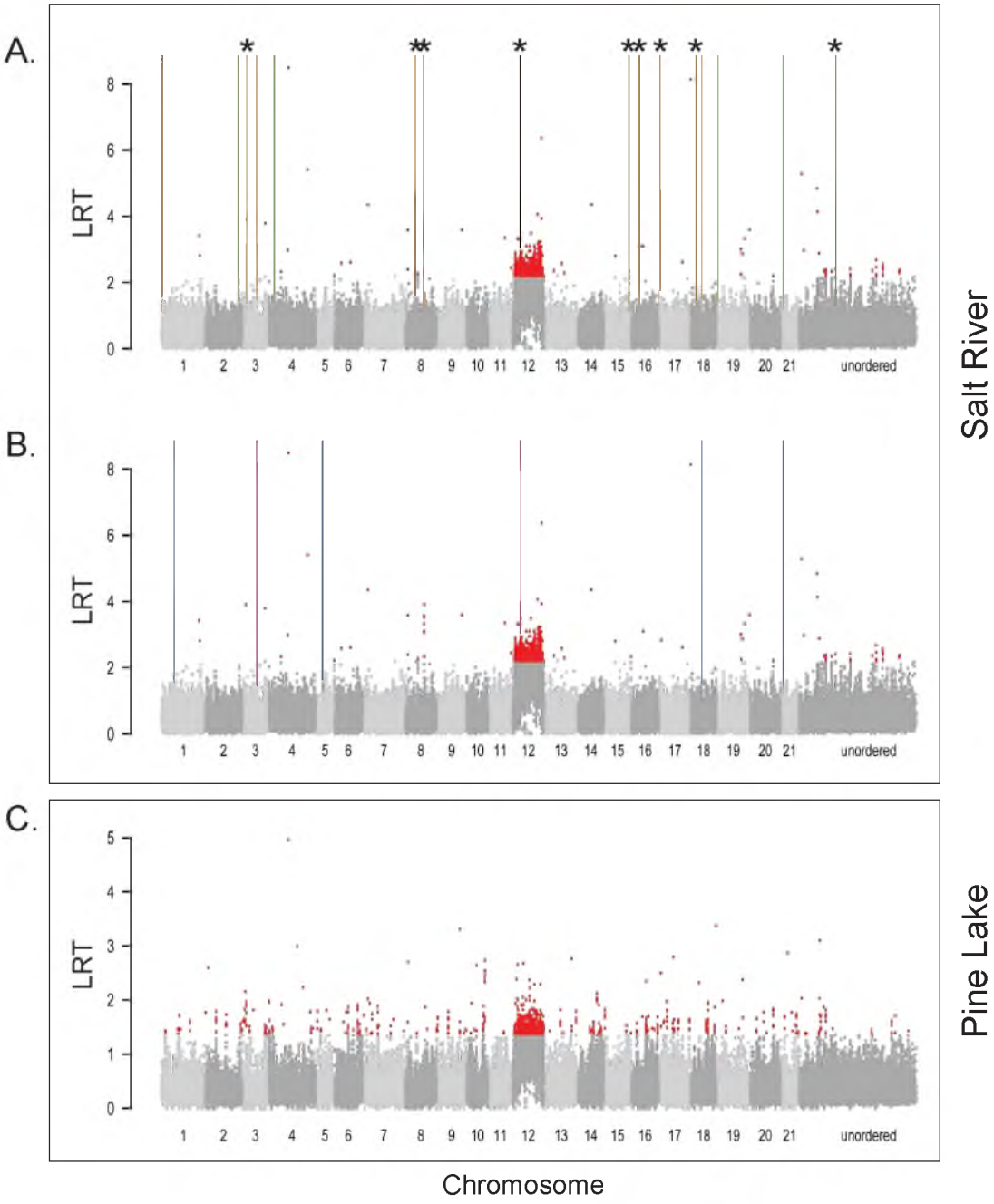


Figure 1.2. STRUCTURE analysis of Salt River and Pine Lake by pelvic phenotype. STRUCTURE plot showing that based on genotypes at 12 unlinked microsatellite markers; Pine Lake and Salt River are distinct populations, but do not show any substructure based on pelvic phenotype within a population.

Figure 1.3. Summary of whole-genome scans and QTL mapping. Likelihood ratio test values averaged in 10kb sliding windows (2kb step) plotted across the genome. Putative linkage groups based on synteny with threespine sticklebacks are pictured from left to right. Any window with an LRT score in the top 0.1% of all windows is indicated by a red point. Quantitative trait loci identified by mapping are indicated by colored vertical lines. A) LRT values and QTL results in Salt River. Colors of vertical lines indicate how many families a given QTL was identified in (green = 1 family, orange = 2 families, black = 3 families). B) LRTvalues and QTL results in Salt River when sexes are analyzed separately. Colors of vertical lines indicate which sex a given QTL was identified in (pink = females only, blue = males only, purple = both sexes). C) LRT values in Pine Lake show broadly similar patterns to those seen in Salt River.


A.
B.
C.
LRT
Salt River
Pine Lake
unordered
Chromosome

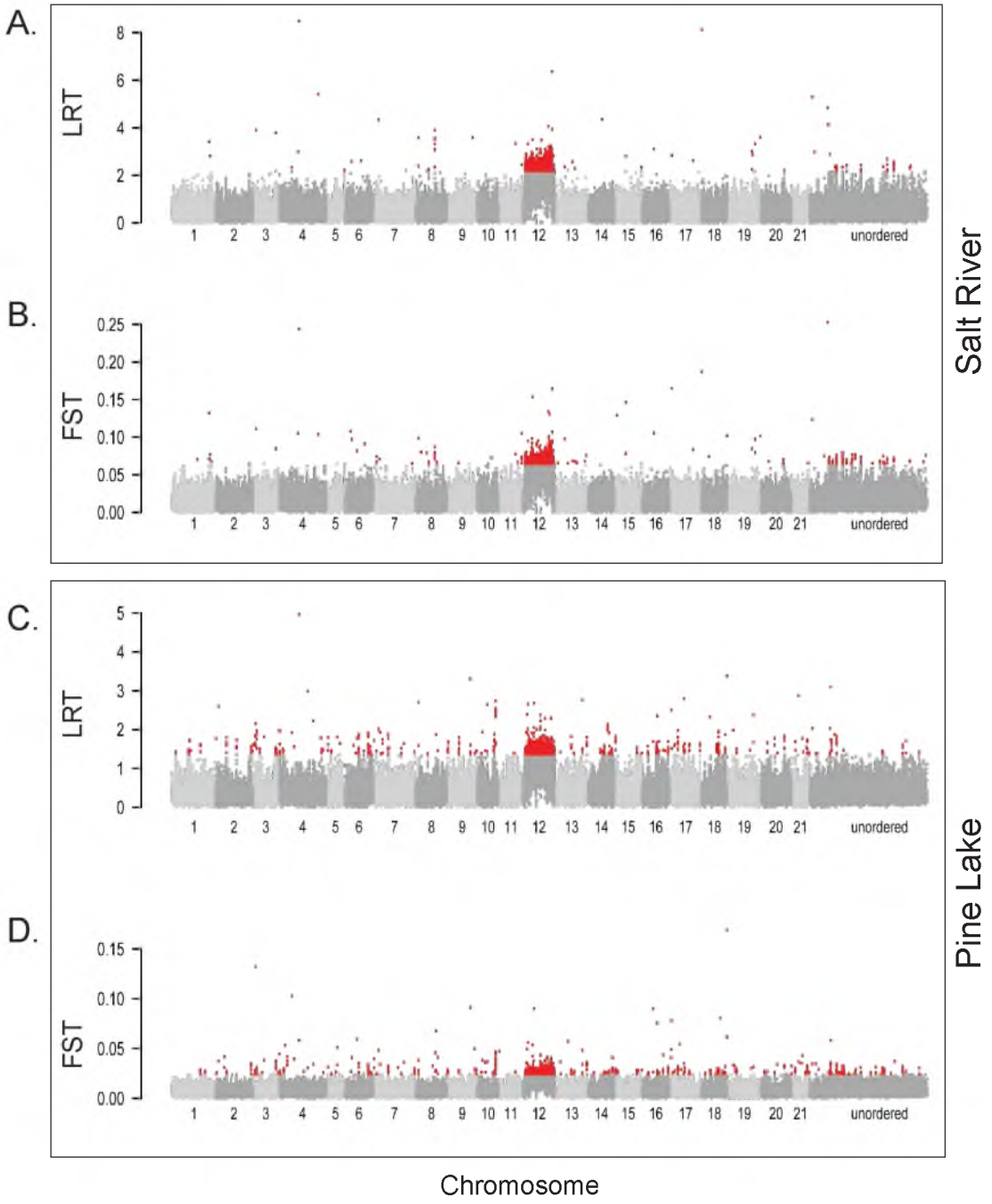


Figure 1.4. Comparison of LRT and F_{ST} from Salt River and Pine Lake.

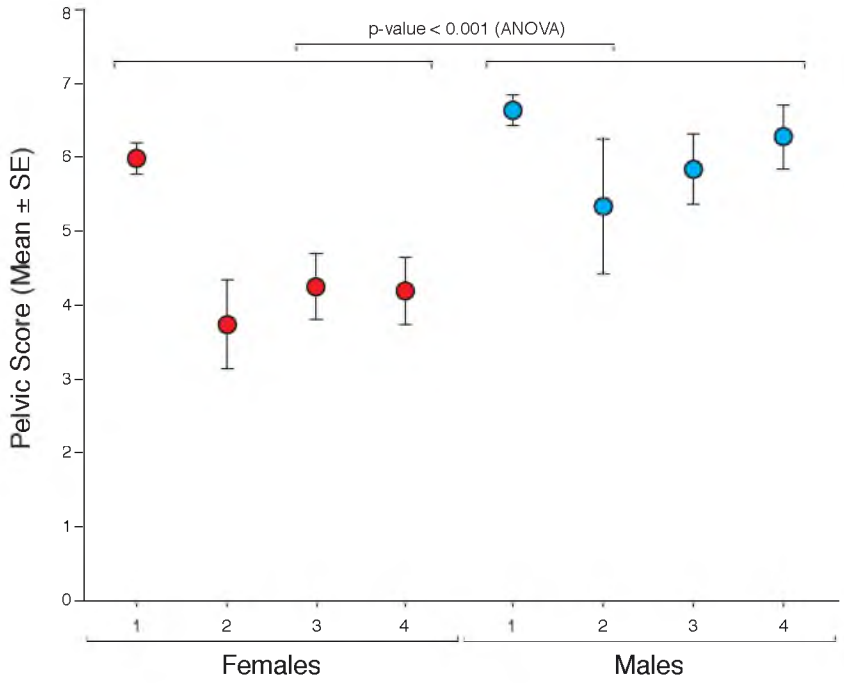


Figure 1.5. Mean pelvic score in Salt River crosses. Mean pelvic score (±SE) separated by family and sex. Females (red) have a significantly lower pelvic score than males (blue). This trend holds when all fish are grouped together (ANOVA; $P < 0.001$) as well as within each family (Family 1, $P < 0.01$; Family 2, $P < 0.01$; Family 3, $P < 0.001$; Family 4, $P < 0.001$).

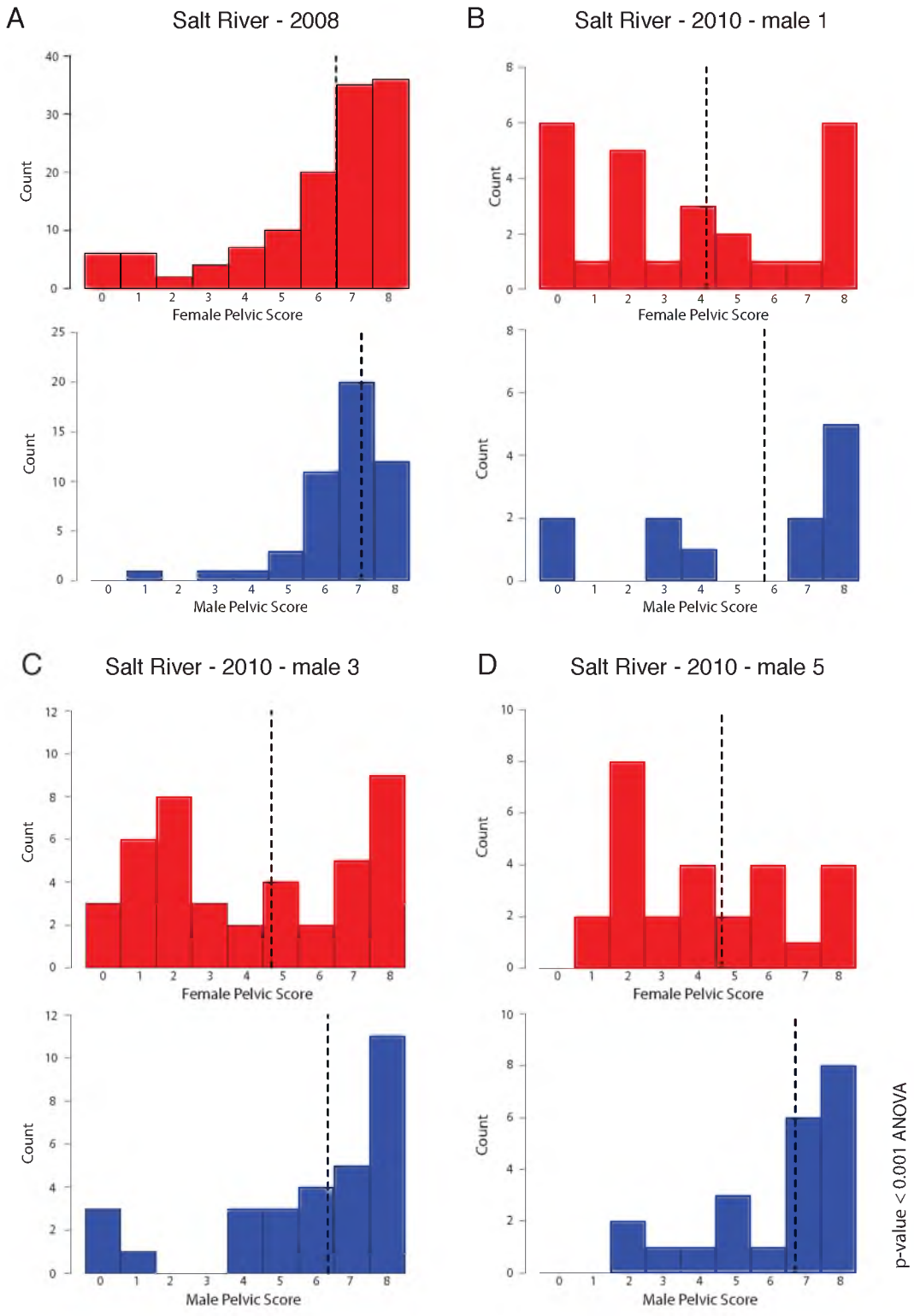
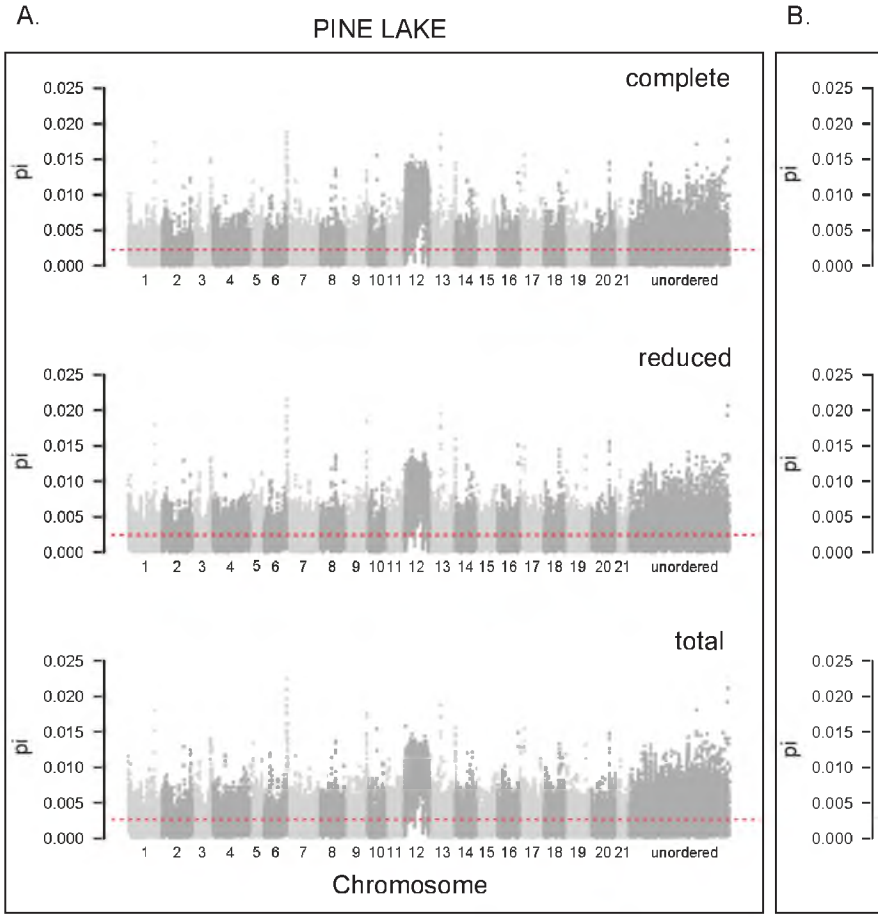
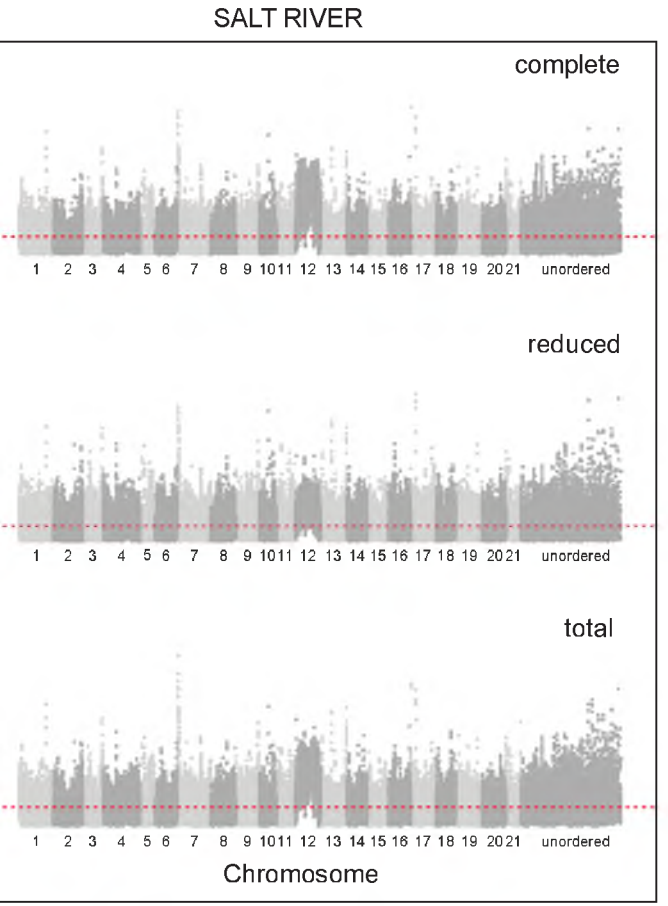


Figure 1.6. Histogram of pelvic scores in Salt River half-sibling families. Female scores are depicted in red and male scores in blue. The mean pelvic phenotype for each group is denoted by a vertical dashed line.

Figure 1.7. Nucleotide diversity (pi) in Pine Lake and Salt River populations. Values of pi were calculated in 10-kb sliding windows with 2-kb steps for complete (top) and reduced (center) pools separately, as well as for the total population (bottom).


A.
PINE LAKE
B.
complete
reduced
total
pi
0.025
0.020
0.015
0.010
0.005
0.000
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 unordered
Chromosome


SALT RIVER
complete
1 2 3 4 5 6 7 8 9 1011 12 13 14 15 16 17 18 19 2021 unordered
reduced
1 2 3 4 5 6 7 8 9 1011 12 13 14 15 16 17 18 19 2021 unordered
total
1 2 3 4 5 6 7 8 9 1011 12 13 14 15 16 17 18 19 2021 unordered
Chromosome

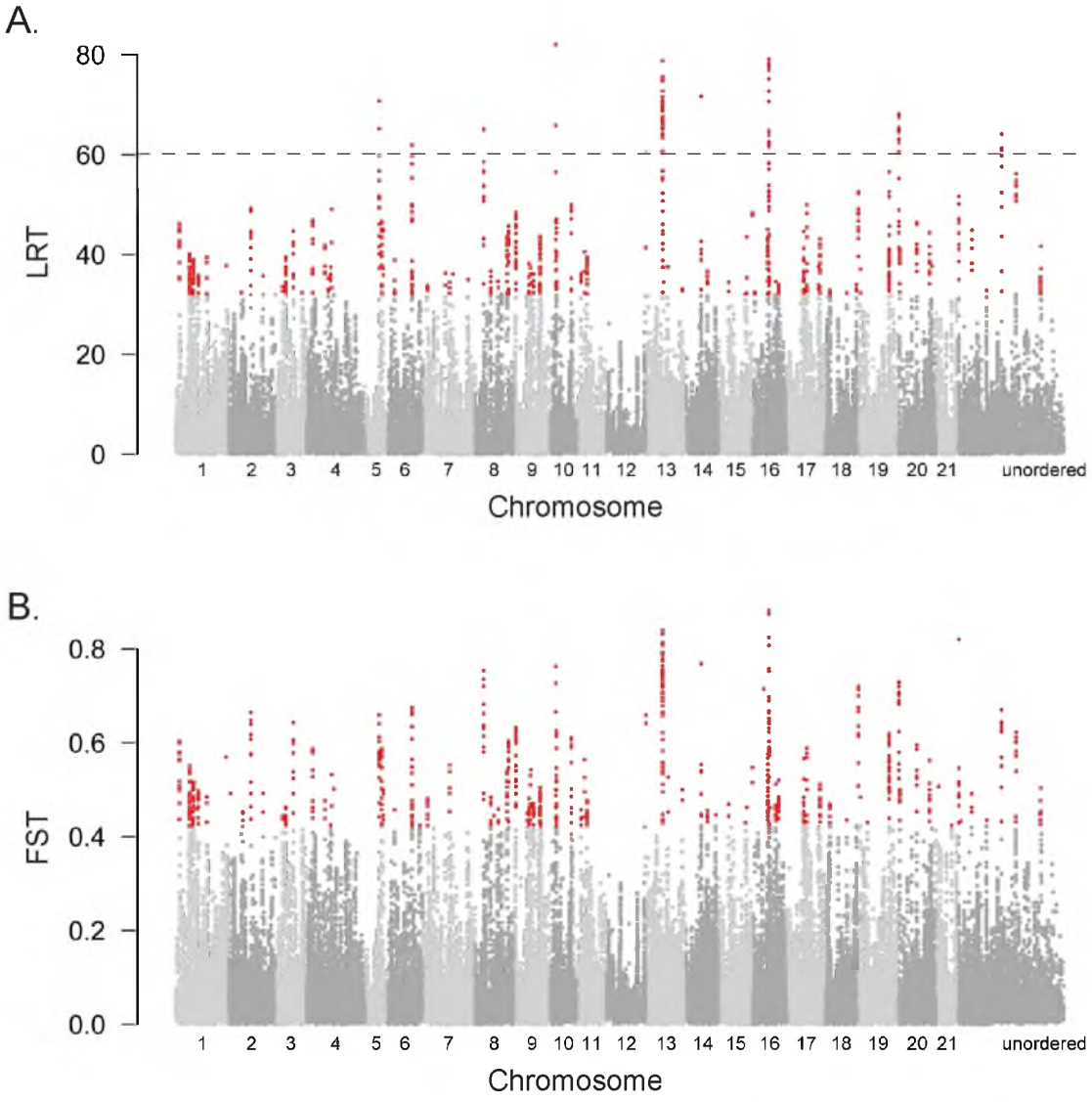


Figure 1.8. LRT and F_{ST} values in interpopulation comparisons. LRT and F_{ST} values in a comparison of Salt River and Pine Lake (averaged in 25-kb sliding windows 2-kb step) plotted across the genome. Putative linkage groups based on synteny with threespine sticklebacks are pictured from left to right. Any window with an LRT score in the top 0.1% of all windows is indicated by a red point.

CHAPTER 2

DIVERGENCE, CONVERGENCE, AND THE ANCESTRY OF FERAL POPULATIONS IN THE DOMESTIC ROCK PIGEON

Reprinted from Curr. Biol., 22, Stringham, S. et al., Divergence, convergence, and the ancestry of feral populations in the domestic rock pigeon, 1–7,

Please cite this article in press as: Stringham et al., Divergence, Convergence, and the Ancestry of Feral Populations in the Domestic Rock Pigeon, Current Biology (2012), doi:10.1016/j.cub.2011.12.045

Current Biology 22, 1–7, February 21, 2012 ©2012 Elsevier Ltd All rights reserved DOI 10.1016/j.cub.2011.12.045

Report

Divergence, Convergence, and the Ancestry of Feral Populations in the Domestic Rock Pigeon

Sydney A. Stringham,[1,3] Elisabeth E. Mulroy,[1,3] Jinchuan Xing,[2] David Record,[1] Michael W. Guernsey,[1] Jaclyn T. Aldenhoven,[1] Edward J. Osborne,[1] and Michael D. Shapiro[1,*]

[1]Department of Biology
[2]Department of Human Genetics
University of Utah, Salt Lake City, UT 84112, USA

Summary

Domestic pigeons are spectacularly diverse and exhibit variation in more traits than any other bird species [1]. In *The Origin of Species*, Charles Darwin repeatedly calls attention to the striking variation among domestic pigeon breeds—generated by thousands of years of artificial selection on a single species by human breeders—as a model for the process of natural divergence among wild populations and species [2]. Darwin proposed a morphology-based classification of domestic pigeon breeds [3], but the relationships among major groups of breeds and their geographic origins remain poorly understood [4, 5]. We used a large, geographically diverse sample of 361 individuals from 70 domestic pigeon breeds and two free-living populations to determine genetic relationships within this species. We found unexpected relationships among phenotypically divergent breeds as well as convergent evolution of derived traits among several breed groups. Our findings also illuminate the geographic origins of breed groups in India and the Middle East and suggest that racing breeds have made substantial contributions to feral pigeon populations.

Results and Discussion

Genetic Structure of Domestic Pigeon Breeds

Charles Darwin was a pigeon aficionado and relied heavily on the dramatic results of artificial selection in domestic pigeons to communicate his theory of natural selection in wild populations and species [2]. "Believing that it is always best to study some special group, I have, after deliberation, taken up domestic pigeons," he wrote in *The Origin of Species* [2] (p. 20). Darwin noted that unique pigeon breeds are so distinct that, based on morphology alone, a taxonomist might be tempted to classify them as completely different genera [3], yet he also concluded that all breeds are simply variants within a single species, the rock pigeon *Columba livia*.

Pigeons were probably domesticated in the Mediterranean region at least 3,000–5,000 years ago, and possibly even earlier as a food source [3, 6, 7]. Their remarkable diversity can be viewed as the outcome of a massive selection experiment. Breeds show dramatic variation in craniofacial structures, color and pattern of plumage pigmentation, feather placement and structure, number and size of axial and appendicular skeletal elements, vocalizations, flight behaviors, and many other traits [1–5]. Furthermore, many of these traits are present in multiple breeds. Today, a large and dedicated pigeon hobbyist community counts thousands of breeders among its ranks worldwide. These hobbyists are the caretakers of a valuable—but largely untapped—reservoir of biological diversity.

Here, as an initial step in developing the pigeon as a model for evolutionary genetics and developmental biology, we address two fundamental questions about the evolution of derived traits in this species. First, what are the genetic relationships among modern pigeon breeds? And second, does genetic evidence support the shared ancestry of breeds with similar traits, or did some traits evolve repeatedly in genetically unrelated breeds?

To address these questions, we studied the genetic structure and phylogenetic relationships among a large sample of domestic pigeon breeds. Our primary goal was to examine relationships among traditional breed groups, to which breeds are assigned based on phenotypic similarities and/or geographic regions of recent breed development (Figure 1) [4, 5, 8]. First, we used 32 unlinked microsatellite markers to genotype 361 individual birds from 70 domestic breeds and two free-living populations. We next used the Bayesian clustering method in STRUCTURE software [9] to detect genetically similar individuals within the sample (Figure 1; see also Figure S1 available online). When two genetic clusters were assumed (K = 2, where K is the number of putative clusters of genetically similar individuals; Figure 1), the first cluster combined several breed groups with dramatically different morphologies. Principal members of this grouping included the pouters and croppers, which have a greatly enlarged, inflatable crop (an outpocketing of the esophagus); the fantails, which have supernumerary and elevated tail feathers; and mane pigeons, breeds with unusual feather manes or hoods about the head (Figure 1).

The second ancestral cluster consisted mainly of the tumblers (including rollers and highflyers), the most breed-rich of the major groups (at least 80 breeds recognized in the USA) [4, 8]. Tumblers are generally small bodied and were originally bred as performance flyers, with many breeds still capable of performing backward somersaults in flight. In most modern tumbler breeds, however, selection is most intense on morphological traits such as beak size and plumage. Also included in this cluster are the owl and the wattle breeds (wattles are skin thickenings emanating from the beak). These two breed groups contrast dramatically in several key traits: owls are typically diminutive in body size, have a pronounced breast or neck frill, and have among the smallest beaks of all breeds, whereas the wattle breeds (English carrier, scandaroon, and dragoon in our analysis) are larger bodied, lack a frill, and have among the most elaborated beak skeletons of all domestic pigeons [4, 5]. The homers (homing pigeons and their relatives) are included in the second cluster as well. The carrier, cumulet, and owl breeds—all members of this cluster—contributed to the modern homing pigeon during its development in England and Belgium approximately 200 years ago [5]. Consistent with this recent admixture, the owls and several homer breeds

[3]These authors contributed equally to this work
*Correspondence: shapiro@biology.utah.edu

Please cite this article in press as: Stringham et al., Divergence, Convergence, and the Ancestry of Feral Populations in the Domestic Rock Pigeon, Current Biology (2012), doi:10.1016/j.cub.2011.12.045

Current Biology Vol 22 No 4
2

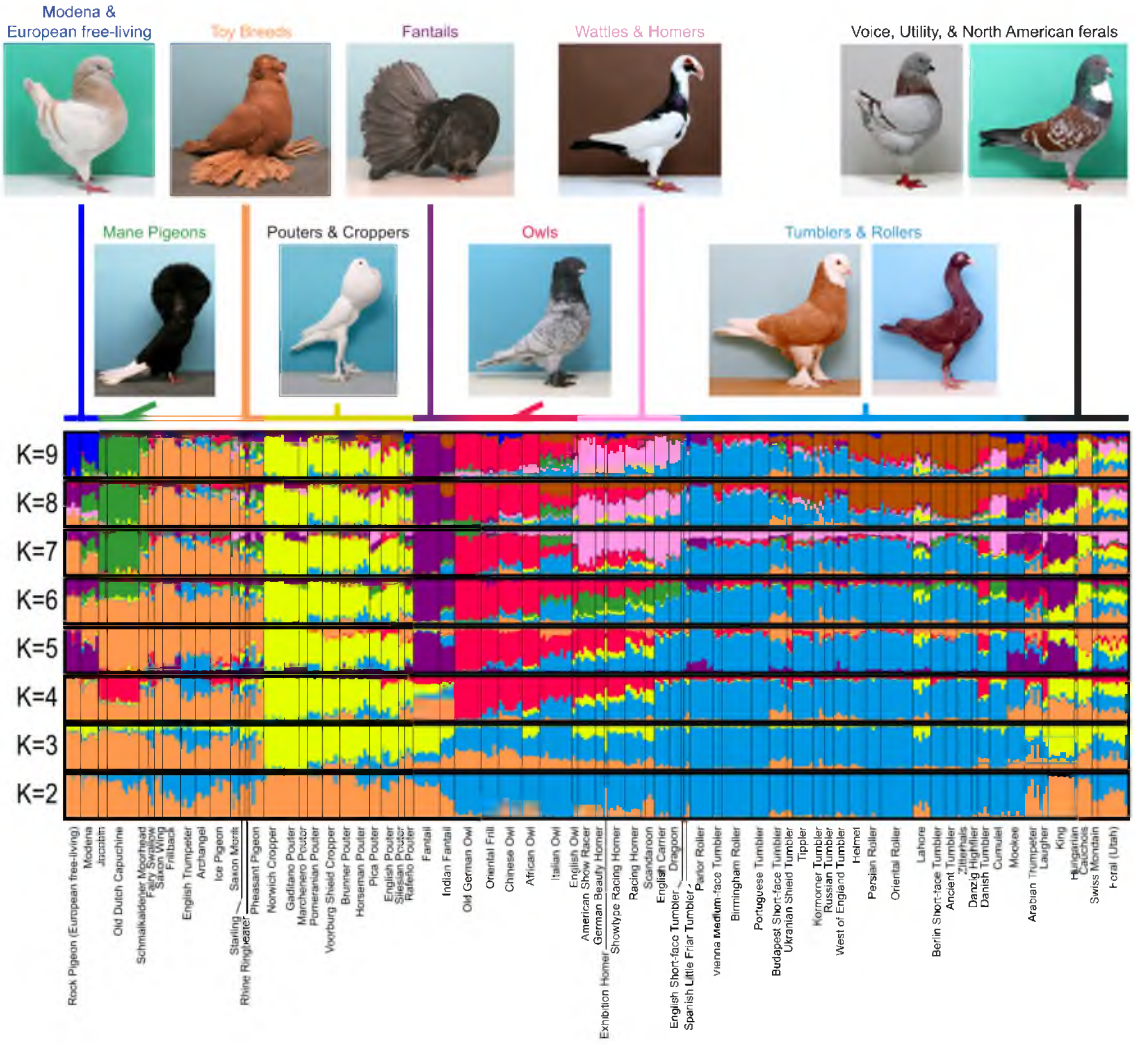


Figure 1. Genetic Structure of the Rock Pigeon (*Columba livia*)

Results from STRUCTURE analysis showing coefficients of genetic cluster membership of 361 individuals representing 70 domestic breeds and two free-living populations (European and North American, at the far left and far right of the plots, respectively) of rock pigeon. Each vertical line represents an individual bird, and proportion of membership in a genetic cluster is represented by different colors. Thin black lines separate breeds. At K = 2, the owls, wattles, and tumblers are the predominant members of one cluster (blue), while other breeds comprise another cluster (orange). At K = 3, the pouters and fantails (yellow) separate from the toys and other breeds, and at K = 5, the fantails separate from the pouters. Pouters and fantails also share genetic similarity with the recently derived king, a breed with a complex hybrid background that probably includes contributions from Indian breeds [5]. At K = 5, fantails are also united with the Modena, an ancient Italian breed, and a free-living European population. The latter two form a discrete cluster at K = 9. At K = 10 and greater (Figure S1), some of the breed groups are assigned to different genetic clusters. This suggests that a number of assumed clusters beyond K = 9 reveals the structure of individual breeds, rather than lending additional insights about genetically similar breed groups. Top row of photos, left to right: Modena, English trumpeter, fantail, scandaroon, king, Cauchois. Bottom row: Jacobin, English pouter, Oriental frill, West of England tumbler, Zitterhals (Stargard shaker). Photos are courtesy of Thomas Hellmann and are not to scale. See Figure S1 for results from K = 2–25 and Tables S1 and S2 for breed and marker information, respectively.

continue to share partial membership in the same cluster at K = 4 and beyond, and the cumulet shares similarity with the homers and wattles at K = 7. Numbers of clusters beyond K = 9 reveal the structure of individual breeds, rather than lending additional insights about breed groups (Figure S1). Notably, although allelic similarity is potentially indicative of shared ancestry, this analysis does not explicitly generate a phylogenetic hypothesis. Moreover, an alternative explanation for clustering is that large effective population sizes might result in an abundance of shared alleles.

We next used multilocus genotype data from a subset of breeds (those with >50% membership in a cluster at K = 9)

Please cite this article in press as: Stringham et al., Divergence, Convergence, and the Ancestry of Feral Populations in the Domestic Rock Pigeon, Current Biology (2012), doi:10.1016/j.cub.2011.12.045

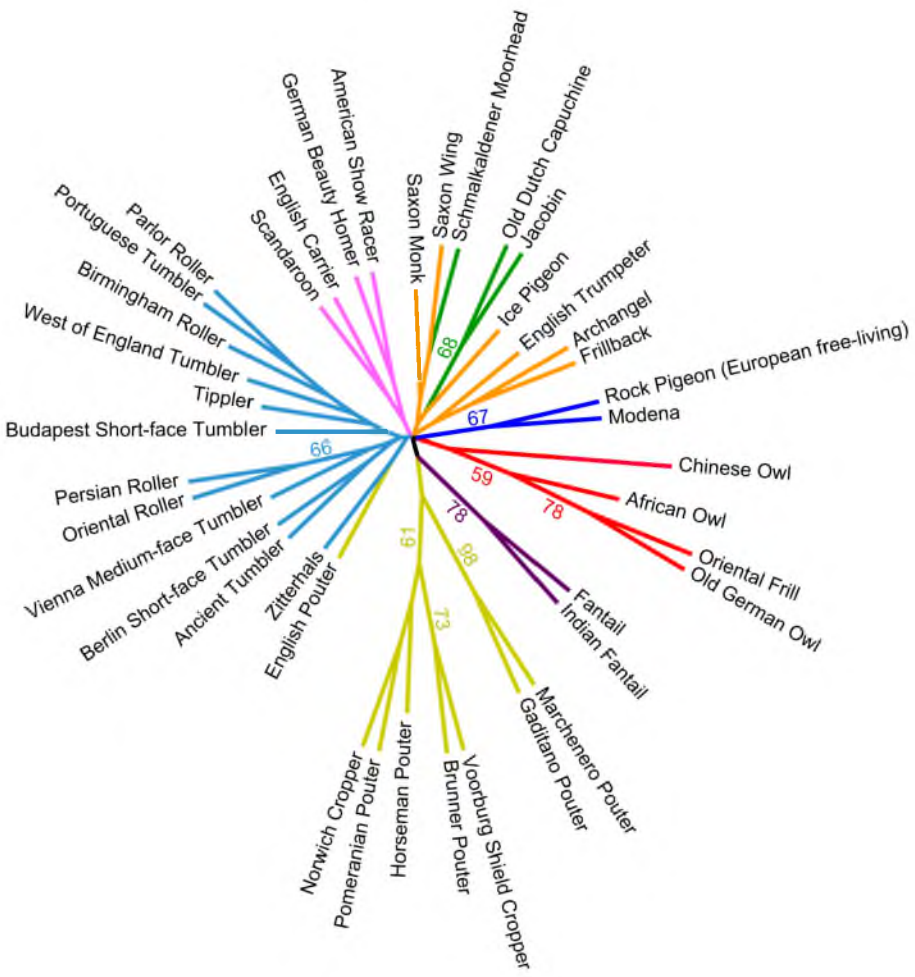


Figure 2. Consensus Neighbor-Joining Tree of Forty Domestic Breeds and One Free-Living Population of Rock Pigeon

The tree here was constructed using pairwise Cavalli-Sforza chord genetic distances and includes the subset of breeds with >50% membership in one genetic cluster at K = 9. Branch colors match cluster colors in Figure 1, except all tumbler breeds are represented with light blue for clarity. A notable incongruence between the STRUCTURE analysis and the tree is the grouping of the English pouter with a tumbler rather than with the other pouters; however, this grouping is not well supported. Percent bootstrap support on branches ($\geq$50%) is based on 1,000 iterations, and branch lengths are proportional to bootstrap values.

to calculate genetic distances among breeds and to generate a neighbor-joining tree (Figure 2). Among the major groups, only subsets of the pouter, fantail, mane, tumbler, Modena and free-living European, and owl branches of the tree have strong statistical support (Figure 2). Nevertheless, at the breed level we observed substantial genetic differentiation, suggesting that in many cases, hybridization among breeds has been limited (mean pairwise F_{ST} = 0.204 for all breeds, maximum F_{ST} = 0.446; potentially more reliable differentiation estimates considering the modest sample sizes for some breeds [10]: mean D_{est} = 0.156, maximum D_{est} = 0.421; Tables S4 and S5). As a comparison, mean pairwise differentiation among African and Eurasian human populations with historically limited gene flow is lower (mean F_{ST} = 0.106, maximum F_{ST} = 0.240 for the comparison between Pygmy and Chinese populations using a dense genome-wide SNP set [11]).

Taking these results together, our analysis shows both expected and unexpected genetic affinities among breeds. Like other domesticated animals such as dogs and chickens, pigeons probably have a reticular rather than hierarchical evolutionary history, which is reflected in the complex genetic structure of many breeds and a star-shaped phylogeny. These findings probably result from hybridization that has occurred throughout the domestication history of the pigeon; this practice continues among some modern breeders as well, often with the goal of transferring a new color into an established breed, or "improving" an existing trait. Unlike the stringent regulations for registering purebred dogs, in which modern breeds are effectively closed breeding populations separated by large genetic distances [12, 13], no barriers exist to mixed ancestry or parentage of pigeons (average F_{ST} = 0.33 between dog breeds [12] compared to 0.24 for pigeons). On the other

Please cite this article in press as: Stringham et al., Divergence, Convergence, and the Ancestry of Feral Populations in the Domestic Rock Pigeon, Current Biology (2012), doi:10.1016/j.cub.2011.12.045

hand, little genetic variation divides dog breeds into subgroups [13], and like our tree (Figure 2), neighbor-joining trees of dogs show limited structuring of the internal branches [12, 13].

Convergent Evolution of Traits

Darwin classified 32 pigeon breeds into four major groups based primarily on morphological traits, especially beak size (Figure 3A). We repeated our STRUCTURE analysis with 14 breeds from Darwin's study that were available to us and found that his morphological classification is broadly congruent with our genetic results (Figure 3B). Beak size is only one of many traits that pigeon breeders have selected over the past several centuries, or in some cases millennia. Feathered feet, head crests, and a multitude of color variants appear in many lineages [8] and must have evolved more than once (Figure 4). Together, these findings suggest that traits do often, but not always, track the ancestry of breeds. This theme of repeated evolution is widespread in genetic studies of other natural and domesticated species as well [14–17].

Geographic Origins of Breeds

Modern breeds are frequently described as having origins in England, Germany, Belgium, or elsewhere in Europe, but their progenitors were probably brought there from afar by traders or colonialists [3–5, 18, 19]. Although we may never definitively know the sites of pigeon domestication, genetic data combined with historical records may provide new clues about the geographic origins of some of the major breed groups.

Most historical accounts trace the origins of the wattle breeds, owls, and tumblers to the Middle and Near East hundreds of years ago, with ancient breeds transported to Europe and India for further development by hybridization or selection [3, 5, 19–21]. Our genetic analyses are consistent with this common geographic origin: these three groups share substantial membership in the same genetic cluster at K = 2–3, and two of the three wattle breeds (English carrier and dragoon) retain high membership coefficients in the tumbler cluster through K = 5 (Figure 1).

The fantail breeds probably originated in India and have undergone less outcrossing than many other breeds [5]. In our STRUCTURE analysis, the fantail (and the Indian fantail to a lesser extent) shows a surprising affinity with the pouters at K = 2–3, and these two groups share a major branch on the neighbor-joining tree (Figures 1 and 2); these two groups are among the most morphologically extreme of all domestic pigeons, and among the most different from each other. European breeders have developed pouters for several hundred years [22, 23], and Dutch traders might have originally brought them to Europe from India [5]. Together, historical accounts and genetic similarity between fantails and pouters support the hypothesis of common geographic origin in India.

Ancestry of Feral Pigeon Populations

Domestic rock pigeons were first brought to North America approximately 400 years ago, and feral populations were probably established shortly thereafter [24, 25]. Likewise, some Eurasian and North African feral populations are probably nearly as old as the most ancient domestication events. In addition to the domestic breeds in our study, we also included a feral pigeon population (Salt Lake City, Utah). Escaped individuals from nearly any domestic breed have the potential to contribute to the feral gene pool, and feral birds showed highly heterogeneous membership across clusters at most values of K (Figure 1). However, we expected that the racing homer would be a major contributor to the feral gene pool. Pigeon racing is an enormously popular and high-stakes hobby worldwide. Although many birds in homing competitions are elite racers that reliably navigate hundreds of miles to their home lofts, some breeders report that up to 20% of their birds that start a race do not return. As predicted, pairwise D_{est} for the racing homer to feral comparison was among the lowest 0.1% of all pairwise comparisons (D_{est} = 0.006), and pairwise F_{ST} was the lowest for any pairwise comparison (F_{ST} = 0.049). Therefore, feral pigeons and racing homers show very little genetic differentiation, and wayward racing homers probably make a substantial contribution to the genetic profile of this local feral population.

We also included samples of free-living rock pigeons (the existence of "pure" wild populations uncontaminated by domestics or ferals is questionable [26]) from Scotland to test for genetic similarities with domestic breeds and with our North American feral sample. Consistent with previous studies [24, 27], European and North American free-living populations are highly differentiated (D_{est} = 0.162). The European sample groups with the Modena, a former racing breed that was developed in Italy up to 2,000 years ago [5] (Figures 1 and 2). This suggests either that Modenas were developed from European free-living populations or that, as in North America, wayward racers contributed to the local feral population, perhaps for centuries. Studies of additional feral populations will reveal whether strong affinities with racing breeds occur locally and sporadically or, as we suspect, almost everywhere.

The Domestic Pigeon as a Model for Avian Genetics and Diversity

Darwin enthusiastically promoted domestic pigeons as a proxy for understanding natural selection in wild populations and species, and pigeons thus hold a unique station in the history of evolutionary biology. More recently, domesticated animals have emerged as important models for rapid evolutionary change [28]. Feathered feet, head ornamentation, skeletal differences, plumage color variation, and other traits prized by breeders offer numerous opportunities to examine the genetic and developmental bases of morphological novelty in birds. These and other traits evolved repeatedly in many breeds, and a challenge arising from this study is to determine whether this distribution of traits resulted from selection on standing variation (either by hybridization between breeds or repeated selection on variants in wild populations), from de novo mutation in independent lineages, or both. In the first case, we would expect certain regions of the pigeon genome to share histories and haplotypes that reflect the transfer of valued traits between breeds. This hypothesis will be testable when we have more detailed information about genomic diversity in this species. Pigeons are also easily bred in the lab, and morphologically distinct breeds are interfertile [2, 3, 29]. Therefore, hybrid crosses should be a fruitful method to map the genetic architecture of derived traits, many of which are known to have a relatively simple genetic basis [4, 29].

The extreme range of variation in domestic pigeons mirrors, if not exceeds, the diversity among wild species of columbids (pigeons and doves) and other birds. Domestic pigeons and

Please cite this article in press as: Stringham et al., Divergence, Convergence, and the Ancestry of Feral Populations in the Domestic Rock Pigeon, Current Biology (2012), doi:10.1016/j.cub.2011.12.045

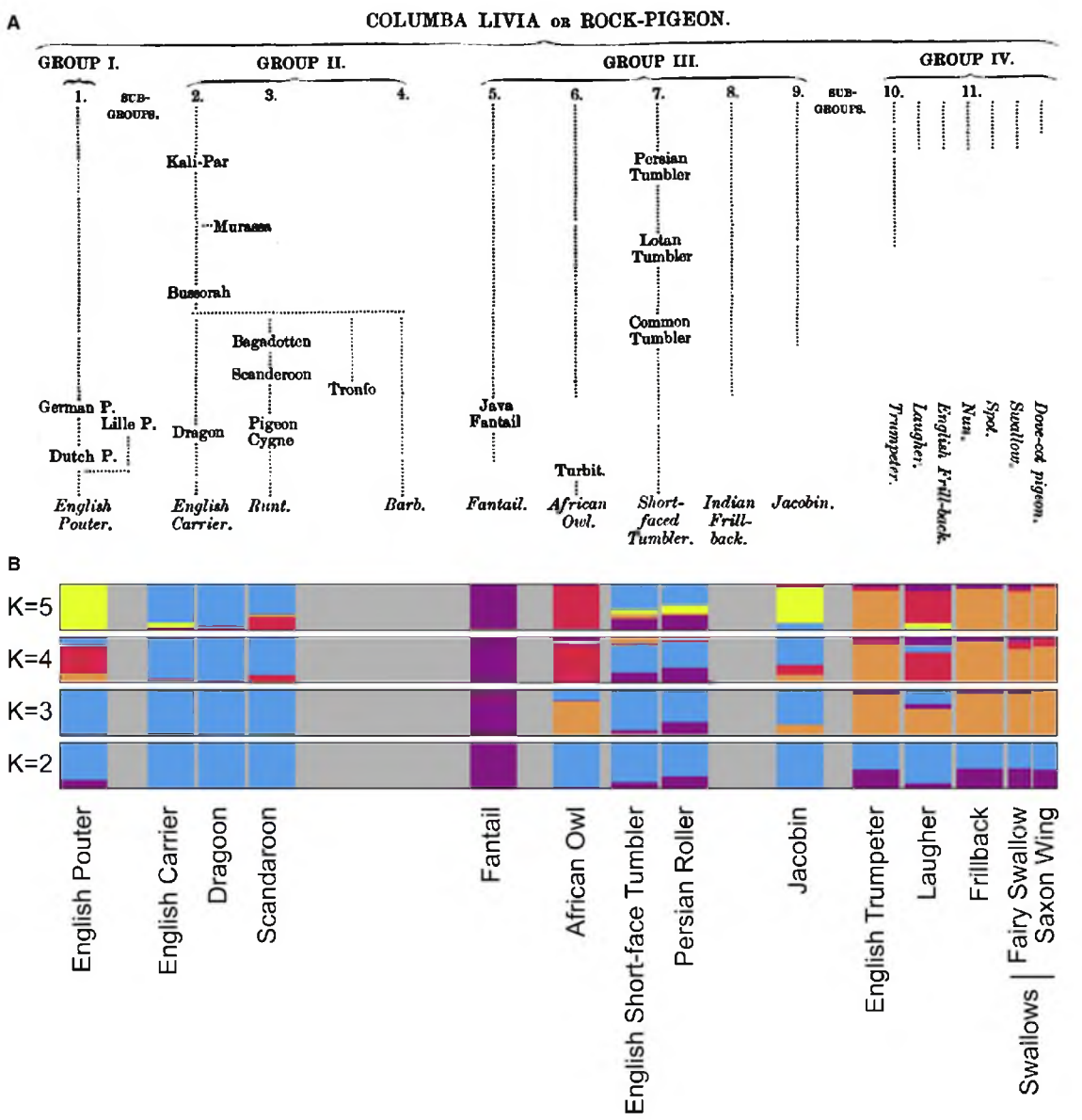


Figure 3. Comparison of Darwin's Morphology-Based Classification and Genetic Structure Analysis of Domestic Pigeon Breeds

(A) Darwin classified 32 breeds into four groups: (I) the pouters and croppers, which have enlarged crops (see also Figures 1 and 4); (II) wattle breeds, many of which have elaborated beaks, and the large-bodied runts; (III) an "artificial" grouping diagnosed by a relatively short beak; and (IV) breeds that resemble the ancestral rock pigeon "in all important points of structure, especially in the beak" [3] (p. 154). Image reproduced with permission from John van Wyhe ed. 2002, The Complete Work of Charles Darwin Online (http://darwin-online.org.uk/).

(B) Mean coefficients of genetic cluster membership for 14 domestic breeds represented in Darwin's classification and our genetic analysis. When two clusters are assumed (K = 2), fantails are separated from all other breeds. At K = 3, the breeds in Darwin's group IV and the African owl (group II) share a high coefficient of membership in a new cluster. At K = 4, the African owl, laugher, and (to a lesser extent) English pouter share membership in a new cluster that includes members of three different morphological groups. At K = 5, the English pouter and Jacobin form a cluster. Although some genetic clusters span more than one morphological group, others are consistent within a group. For example, the wattle breeds (group II), tumblers (group III), and most of group IV remain united with breeds of similar morphology at K = 2–5. Taken together, these results confirm that morphology is a good general predictor of genetic similarity in domestic pigeons, yet they also show that breeds that share allelic similarity can be morphologically distinct. Darwin, too, recognized that breeds united in form were not necessarily united in ancestry and, conversely, that anatomically dissimilar breeds might be related. For example, he classified the short-beaked barb (not in our genetic data set) with the long-beaked breeds of group II.

Please cite this article in press as: Stringham et al., Divergence, Convergence, and the Ancestry of Feral Populations in the Domestic Rock Pigeon, Current Biology (2012), doi:10.1016/j.cub.2011.12.045

Current Biology Vol 22 No 4
6

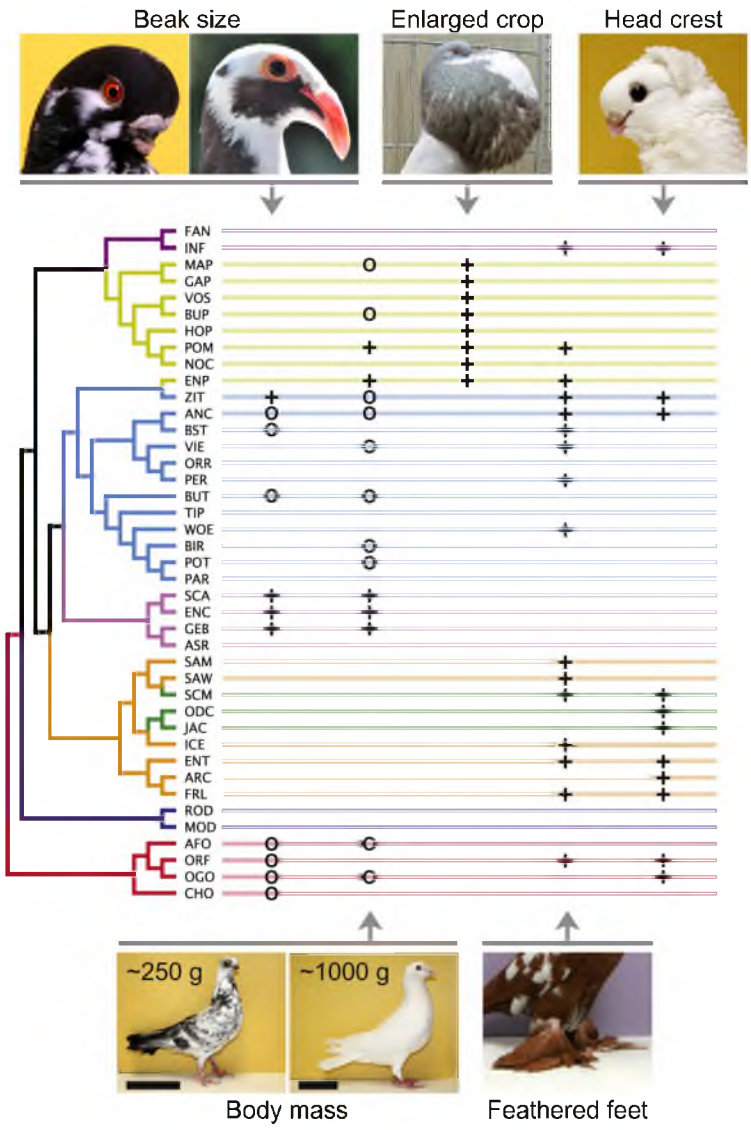


Figure 4. Distribution of Several Derived Traits across Groups of Domestic Pigeons

The phylogenetic tree in Figure 2 was converted to a cladogram format with equal branch lengths (far left). For the beak size column, "+" indicates a substantial increase in size relative to the ancestral condition, and "O" indicates a decrease [4, 8]. For body mass, "+" indicates breeds with a maximum over 550 g, and "O" indicates those under 340 g [4, 8]. Although a 4-fold difference in body mass is depicted here, extremes in body mass among all known breeds differ by more than an order of magnitude. For crop, feathered feet, and head crest, "+" indicates fixed or variable presence of the trait (substantial departure from the ancestral condition [4, 8]). All traits shown were selected in multiple groups except an enlarged crop, which is confined to the pouters and croppers. A possible exception is the Cauchois (not included in the tree; see Figure 1), a non-pouter breed with an enlarged and inflatable crop, thought to have been developed centuries ago from a cross between a pouter and large-bodied Mondain breed [5, 33]. Our STRUCTURE analysis supports this hypothesis, with the Cauchois sharing 37.8%–89.7% membership in the genetic cluster containing the pouters at K = 2–9 (Figure 1). Breeds shown (clockwise from upper left) are African owl*, scandaroon, Norwich cropper, old German owl, West of England tumbler*, white Carneau, and Budapest short-face tumbler. Scale bars represent 10 cm. *Photos courtesy of Thomas Hellmann.

wild bird species vary in many of the same traits, so domestic pigeons provide an entry point to the genetic basis of avian evolutionary diversity in general [1, 30]. Changes in the same genes, and even in some cases the same mutations, have recently been shown to underlie similar phenotypes in both wild and domesticated populations [31, 32]. The genetic history of pigeons is a critical framework for the analysis of the genetic control of many novel traits in this fascinating avian species.

Accession Numbers

The microsatellite markers and sequences reported in this paper have been deposited at GenBank with the accession numbers GF111523–GF111539.

Supplemental Information

Supplemental Information includes one figure, four tables, and Supplemental Experimental Procedures and can be found with this article online at doi:10.1016/j.cub.2011.12.045.

Acknowledgments

We thank Kyle Christensen and members of the Utah Pigeon Club, National Pigeon Association, and Bund Deutscher Rassegflügelzüchter for their spirited collaboration; Elena Boer, Terry Dial, Jennifer Koop, Matt Miller, and Jessica Waite for collection assistance; Jon Seger, Kyle Christensen, and Eric Domyan for comments on drafts of the manuscript; and Thomas Hellmann for photos used in Figures 1 and 4. Animal protocols were approved by the University of Utah Institutional Animal Care and Use Committee (protocol 09-04015). This work was supported by National Institutes of Health (NIH) grant T32GM007464 (S.A.S. and E.J.O.), National

Please cite this article in press as: Stringham et al., Divergence, Convergence, and the Ancestry of Feral Populations in the Domestic Rock Pigeon, Current Biology (2012), doi:10.1016/j.cub.2011.12.045

Science Foundation grant DGE0841233 (S.A.S.), the University of Utah BioURP and UROP programs (E.E.M. and M.W.G.), NIH/National Human Genome Research Institute grant K99HG005846 (J.X.), a Burroughs Wellcome Fund Career Award in the Biomedical Sciences (M.D.S.), and a gift from Onorio Catenacci.

Received: September 2, 2011
Revised: December 19, 2011
Accepted: December 19, 2011
Published online: January 19, 2012

References

1. Price, T.D. (2002). Domesticated birds as a model for the genetics of speciation by sexual selection. Genetica *116*, 311–327.
2. Darwin, C. (1859). On the Origin of Species by Means of Natural Selection (London: John Murray).
3. Darwin, C.R. (1868). The Variation of Animals and Plants under Domestication, *Volume 1* (London: John Murray).
4. Levi, W.M. (1965). Encyclopedia of Pigeon Breeds (Sumter, SC: Levi Publishing).
5. Levi, W.M. (1986). The Pigeon, Second Revised Edition (Sumter, SC: Levi Publishing).
6. Sossinka, R. (1982). Domestication in birds. In Avian Biology, *Volume 6*, D.S. Farner, A.S. King, and K.C. Parkes, eds. (London: Academic Press), pp. 373–403.
7. Driscoll, C.A., Macdonald, D.W., and O'Brien, S.J. (2009). From wild animals to domestic pets, an evolutionary view of domestication. Proc. Natl. Acad. Sci. USA *106* (*Suppl 1*), 9971–9978.
8. National Pigeon Association. (2010). National Pigeon Association Book of Standards (Goodlettsville, TN: Purebred Pigeon Publishing).
9. Pritchard, J.K., Stephens, M., and Donnelly, P. (2000). Inference of population structure using multilocus genotype data. Genetics *155*, 945–959.
10. Jost, L. (2008). G_{ST} and its relatives do not measure differentiation. Mol. Ecol. *17*, 4015–4026.
11. Xing, J., Watkins, W.S., Witherspoon, D.J., Zhang, Y., Guthery, S.L., Thara, R., Mowry, B.J., Bulayeva, K., Weiss, R.B., and Jorde, L.B. (2009). Fine-scaled human genetic structure revealed by SNP microarrays. Genome Res. *19*, 815–825.
12. Parker, H.G., Kim, L.V., Sutter, N.B., Carlson, S., Lorentzen, T.D., Malek, T.B., Johnson, G.S., DeFrance, H.B., Ostrander, E.A., and Kruglyak, L. (2004). Genetic structure of the purebred domestic dog. Science *304*, 1160–1164.
13. Vaysse, A., Ratnakumar, A., Derrien, T., Axelsson, E., Rosengren Pielberg, G., Sigurdsson, S., Fall, T., Seppälä, E.H., Hansen, M.S., Lawley, C.T., et al.; LUPA Consortium. (2011). Identification of genomic regions associated with phenotypic variation between dog breeds using selection mapping. PLoS Genet. *7*, e1002316.
14. Aldenhoven, J.T., Miller, M.A., Corneli, P.S., and Shapiro, M.D. (2010). Phylogeography of ninespine sticklebacks (*Pungitius pungitius*) in North America: glacial refugia and the origins of adaptive traits. Mol. Ecol. *19*, 4061–4076.
15. Colosimo, P.F., Hosemann, K.E., Balabhadra, S., Villarreal, G., Jr., Dickson, M., Grimwood, J., Schmutz, J., Myers, R.M., Schluter, D., and Kingsley, D.M. (2005). Widespread parallel evolution in sticklebacks by repeated fixation of Ectodysplasin alleles. Science *307*, 1928–1933.
16. Arendt, J., and Reznick, D. (2008). Convergence and parallelism reconsidered: what have we learned about the genetics of adaptation? Trends Ecol. Evol. (Amst.) *23*, 26–32.
17. Hovav, R., Chaudhary, B., Udall, J.A., Flagel, L., and Wendel, J.F. (2008). Parallel domestication, convergent evolution and duplicated gene recruitment in allopolyploid cotton. Genetics *179*, 1725–1733.
18. Lyell, J.C. (1881). Fancy Pigeons (London: A. Bradley).
19. Tegetmeier, W.B. (1868). Pigeons: Their Structure, Varieties, Habits, and Management (London: George Routledge and Sons).
20. Alla-oodeen (1888). The art of training pigeons in the East. The Zoologist (Lond.) *12*, 209–219, 252–258.
21. Fazl, A. (1888). The art of training pigeons in the East. The Zoologist (Lond.) *12*, 167–174.
22. Aldrovandi, U. (1610). Ornithologiae (Frankfurt: Nicolai Bassaei).
23. Ray, J. (1676). The Ornithology of Francis Willughby (London: John Martin).
24. Johnston, R.F. (1994). Geographic variation of size in feral pigeons. Auk *111*, 398–404.
25. Schorger, A.W. (1952). Introduction of the domestic pigeon. Auk *69*, 462–463.
26. Goodwin, D. (1983). Pigeons and Doves of the World, Third Edition (Ithaca, NY: Comstock Publishing Associates).
27. Johnston, R.F., Siegel-Causey, D., and Johnson, S.G. (1988). European populations of the rock dove *Columba livia* and genotypic extinction. Am. Midland Nat. *120*, 1–10.
28. Akey, J.M., Ruhe, A.L., Akey, D.T., Wong, A.K., Connelly, C.F., Madeoy, J., Nicholas, T.J., and Neff, M.W. (2010). Tracking footprints of artificial selection in the dog genome. Proc. Natl. Acad. Sci. USA *107*, 1160–1165.
29. Sell, A. (1994). Breeding and Inheritance in Pigeons (Hengersberg, Germany: Schober Verlags-GmbH).
30. Baptista, L.F., Gomez Martinez, J.E., and Horblit, H.M. (2009). Darwin's pigeons and the evolution of the columbiforms: recapitulation of ancient genes. Acta Zool. Mex. *25*, 719–741.
31. Arnaud, N., Lawrenson, T., Østergaard, L., and Sablowski, R. (2011). The same regulatory point mutation changed seed-dispersal structures in evolution and domestication. Curr. Biol. *21*, 1215–1219.
32. Hoekstra, H.E. (2006). Genetics, development and evolution of adaptive pigmentation in vertebrates. Heredity (Edinb.) *97*, 222–234.
33. Buffon, G. (1774). Le Pigeon. In Histoire Naturelle, *Volume 43* (Paris: Imprimerie de F. Dufart), pp. 154–311.

CHAPTER 3

THE GENETIC BASIS OF DIVERGENCE AND CONVERGENCE IN TELEOST FISH

Abstract

Despite longstanding interest in how vertebrates acquire novel traits that characterize macroevolutionary transformations, remarkably little is known about the number, location, and types of mutations that control major differences among vertebrate lineages. As major transformations among vertebrates occurred in the distant past, traditional genetic approaches typically will not work to understand the genes that actually mattered in the evolution of key innovations. This is because extant taxa with disparate traits or body plans are usually too distantly related for experimental genetics. In a limited number of extant species, however, different populations have evolved anatomical, physiological, or behavioral changes of a magnitude that typically characterizes different species. Not many species meet this criterion, but the ones that do are emerging as key models in evolutionary genetics and developmental biology. Using these special cases, we can gain important insights about the genetic architecture of adaptive traits, the types of mutations that occur (and their developmental consequences), and whether similar kinds of mutations occur repeatedly when similar traits evolve independently. In this chapter, we address these themes in the context of three diverse lineages of teleost fishes: sticklebacks, Mexican cave tetras, and African cichlids.

> "If it could be demonstrated that any complex organ existed, which could not possibly have been formed by numerous, successive, slight modifications, my

theory would absolutely break down." Charles Darwin (1859, p.189)

"… Nature does make jumps now and then, and a recognition of the fact is of no small importance in disposing of many minor objections to the doctrine of transmutation."
Thomas Huxley (1860, p.310)

Introduction

Great transformations among the vertebrates can only be appreciated and understood by elucidating the microtransformational mechanisms responsible for form and function. However, when studying major transformations that occurred many millions of years ago, we have limited access to the molecular mechanisms underlying these changes. For example, evolutionary biologists can only dream of using controlled genetic crosses between birds and nonavian theropod dinosaurs to map the key genetic changes in the evolution of flight or crossing a fish and a tetrapod to identify the genes that matter in fin versus limb development and function. Even among extant vertebrates, anatomically divergent species are typically too distantly related to allow traditional genetic approaches, which require the production of fertile offspring. Moreover, although the complete sequences of many vertebrate genomes are now available, determining which of the millions of DNA sequence and structural differences among species are actually responsible for particular trait differences remains a major challenge.

Organismal diversity, and morphological diversity in particular, is rooted in changes to developmental programs. That is, major anatomical changes among adults of different populations and species must manifest sometime between fertilization of an egg and sexual maturity. Developmental differences, in turn, are regulated largely (but by no means exclusively) by changes in genetic programs. Much of what we know about the molecular genetic basis of vertebrate development comes from mechanistic studies of

traditional laboratory models such as the mouse, chicken, African claw-toed frog, and zebrafish. Despite major advances in our understanding of organismal construction from normal and mutant inbred laboratory populations, we know considerably less about the genetic and developmental basis of *natural variation* among vertebrates. Evolutionary developmental genetics (often referred to as "evo-devo") takes advantage of variation in the wild to directly address the link between genotype and phenotype among species, which will lead to a better understanding of the molecular origins of diversity.

In contrast to most other chapters in this volume, we focus on variation and transformations among populations and closely related species. This scale of investigation has the advantage of using traditional genetic approaches to understand vertebrate diversity, a strategy that typically is not available when studying major transformations among lineages with distant common ancestors. By understanding the genetic changes that underlie phenotypic changes at the micro-evolutionary scale, we can begin to address central questions about the mechanisms underlying morphological transformations within and among species. For example, how many genetic changes underlie substantial morphological changes? Where do these changes occur, in the coding or regulatory regions of genes? Finally, do the same genetic changes underlie the repeated evolution of similar traits in different populations and species? This framework is generally limited to addressing questions about diversity between groups that can be interbred or share close genomic similarity (typically populations within a species). However, results of these studies also provide clues to macroevolutionary patterns, as similar molecular mechanisms are sometimes involved in the evolution of similar traits in distantly related organisms.

We focus here on examples of particularly striking variation in teleost fishes, With nearly 29,000 extant species (Santini et al. 2009), teleosts are among the most successful radiations of vertebrates. In some cases, changes among populations *within* a species are so pronounced that they resemble in magnitude the differences *among* species. These cases of intraspecific variation in extant taxa are especially important to our understanding of the mechanisms that give rise to phenotypic transformations and perhaps ultimately to new species and adaptive radiations. Within teleosts, we discuss examples of genetic mechanisms of diversification in sticklebacks, Mexican cavefish, and African cichlids. Each of these groups evolved dramatic—and repeated—phenotypic transformations in response to novel habitats, and each provides an ideal framework to examine the genetic basis of organismal diversity. These are not the only teleost groups in which the genetic basis of variation has been studied; however, the traits and transformations we highlight below introduce important themes and trends in the evolution of teleosts and other vertebrates.

Each of these groups of teleosts also offers important advantages as a model system in evolutionary genetics. First, different populations or closely related species within each group can be interbred to produce fertile offspring. This important characteristic facilitates traditional genetic mapping of traits of interest. Second, all three groups have been studied for many decades from the perspectives of ecology, natural history, and to a lesser extent, classical genetics and developmental biology. This foundation provides an important entry point to dissect the molecular genetic changes that control organismal diversity. Below, we consider micro-evolutionary transformations in each group then discuss their impact on our understanding of broader trends of the

genetic basis of vertebrate diversity.

Sticklebacks (Family Gasterosteidae)

Introduction

Sticklebacks comprise seven species of small teleost fish that are widespread and often locally abundant across the Northern Hemisphere. A subset of these species exhibits tremendous intraspecific variation in skeletal morphology, body shape, color, behavior, and physiological adaptations. The most recent adaptive radiation of the threespine stickleback began with the retreat of glacial ice less than 20,000 years ago (Bernatchez and Wilson 1998; Hewitt 2000). This retreat created new inland freshwater habitats, which were subsequently colonized by marine stickleback populations. The transition to resident freshwater environments presented novel trophic, predatory, and physiological challenges. For example, freshwater habitats vary dramatically from marine habitats in temperature, topological complexity, water chemistry, and predator loads (Heuts 1947; Hagen and Gilbertson 1973b; Moodie et al. 1973; Hagen and Moodie 1982; Coad 1983; Giles 1983; Reimchen 1992, 1995; Kitano et al. 2008).

Geographically and phylogenetically distant populations of threespine sticklebacks have evolved strikingly similar suites of characteristics in response to the shift to freshwater habitats. For example, many populations have lost major components of their bony armor, including the lateral plates and pelvic girdle, in response to new predator loads and other factors (Bell and Foster 1994) (Figure 3.1). Furthermore, parallel phenotypic changes occur not only among populations of threespine sticklebacks (*Gasterosteus aculeatus*), the focus of most recent genetic and genomic studies, but also across species that diverged millions of years ago (e.g., the ninespine stickleback

Pungitius pungitius and the brook stickleback *Culaea inconstans*) (Nelson and Atton 1971; Wootton 1976; Blouw and Boyd 1992; Bell and Foster 1994; Ziuganov and Zotin 1995). Thus, this multispecies system provides an excellent model to examine the genetics of adaptive traits on both micro- and macroevolutionary levels.

Armor Plate Variation

Armor plates are composed of thin dermal bone and almost completely cover the lateral sides of marine threespine sticklebacks ("complete morph"; Figure 3.1a top). In contrast, the number and size of these plates is reduced in most freshwater populations ("low morph"; Figure 3.1a bottom) in response to strong selection in freshwater habitats (discussed below), and the genetic basis of this variation has been the subject of classical genetic studies for decades (Hagen and Gilbertson 1973a; Avise 1976; Ziuganov 1983; Banbura 1994). Laboratory crosses between different morphs showed that probably only a few genes control most of the variation in plate number (Hagen and Gilbertson 1973a; Avise 1976; Ziuganov 1983; Banbura 1994).

More recently, Colosimo et al. (2004) used a molecular genetic approach to identify the major locus controlling plate reduction. To do this, they crossed a complete morph marine fish (Hokkaido Island, Japan) to a low morph freshwater fish (Paxton Lake, British Columbia); the grandchildren (F_2 progeny) of this cross showed a wide range of plate morphologies, including fish that had high or low numbers of plates like their grandparents. By looking for associations between plate phenotypes and segments of chromosomes inherited from either the complete- or low-morph grandparent, Colosimo et al. (2004) found a single position in the genome (a quantitative trait locus, or QTL) on linkage group (LG) 4 that largely determined whether fish had the complete, partial, or

low plate morph (see Figure 3.2). Other studies suggested that LG4 controls plate phenotypes in multiple populations of threespine sticklebacks (Cresko et al. 2004; Schluter et al. 2004). However, key questions remained: which gene(s) in the major QTL region controlled armor variation, and were the mutations the same or different among the many populations with low plates?

Further genetic mapping studies showed that variation in the gene *Ectodysplasin* (*Eda*) was the most likely cause of armor diversity (Colosimo et al. 2005). In vertebrates, *Eda* plays a key role in the development of several tissues derived from the ectoderm, including hair, teeth, sweat glands, and scales (Thesleff and Mikkola 2002; Kangas et al. 2004; Harris et al. 2008). The external armor of sticklebacks is also derived from ectoderm. Importantly, Colosimo et al. (2005) showed that by injecting low-plated embryos with an engineered DNA construct containing a functional version of *Eda*, they could partially restore plate formation in low-plated fish. This provided functional evidence that *Eda* plays a critical role in plate development.

Strikingly, nearly every low-plated population throughout the range of the species appears to have the *same* chromosome segment containing the *Eda* gene (Colosimo et al. 2005). This indicates that the repeated evolution of low plates probably resulted from selection on the same mutant version of *Eda*, rather than by independent mutations in *Eda* in each population. The key to the spread of the low plate allele resides in the marine populations that colonize new freshwater habitats: the low plate version of *Eda* typically found in freshwater populations is also found in a small proportion of marine fish, suggesting that high-plated ocean populations are a "genetic reservoir" for the low plate allele (Colosimo et al. 2005). Once the allele enters a freshwater habitat with the arrival

new marine colonists, selection drives it to high frequency. Transition from high to low plates can happen very quickly. In one Alaskan lake population, for example, Bell et al. (2004) observed a dramatic shift from predominantly high-plated to low-plated in less than 12 years (also see Kitano et al. 2008). Paradoxically, while the genetic basis for this trait is well understood and there is strong evidence for selection on plate phenotypes and the *Eda* locus, the ecological mechanism driving selection is less clear (reviewed in Barrett 2010).

Reduction and Loss of the Pelvic Fin Complex

In addition to variation in lateral armor, at least 20 freshwater populations of threespine stickleback also exhibit reduction or loss of the pelvis (Bell 1974a; Moodie 1976; Campbell and Williamson 1979; Edge and Coad 1983; Bell 1987). The stickleback pelvis is homologous to the pelvic fin skeleton of other teleosts as well as to the tetrapod hindlimb. It is composed of a pelvic girdle and serrated pelvic spines that provide protection from gape-limited predators such as large piscivorous fish (Hoogland et al. 1957; Hagen and Gilbertson 1972; Moodie 1972; Gross 1978; Lescak and von Hippel 2011) (Figure 3.1, B–D). However, reduction of pelvic structures is advantageous in some populations where grasping predators such as aquatic invertebrates are a greater threat, especially to juvenile fish (Hoogland et al. 1957; Reimchen 1980, 1983; Bell et al. 1993; Bell and Orti 1994; Bourgeois et al. 1994). Large pelvic skeletons could be disadvantageous in these habitats because spines provide an additional surface for insects to capture and hold their prey (Reimchen 1980; Reist 1980; Ziuganov and Zotin 1995; Marchinko 2009).

Using a QTL mapping approach similar to the armor plate study, Shapiro et al. (2004) identified the gene *Pitx1* as a major influence on pelvic morphology. In other

vertebrates, *Pitx1* contributes to hindlimb identity and development: mice with an inactive form of the gene exhibit reduced and malformed hindlimbs but normal forelimbs (Lanctôt et al. 1999; Marcil et al. 2003). Furthermore, sticklebacks from the cross that retained pelvic spines showed a marked asymmetry with larger spines on the left side, a feature also seen in the limbs of mice with a nonfunctional (knockout) version of *Pitx1* and humans with a *Pitx1* mutation (Lanctôt et al. 1999; Gurnett et al. 2008).

Unlike in the mouse *Pitx1* knockout, mutations were not found in the coding region of *Pitx1* in pelvisless freshwater stickleback populations compared to marine fish (Shapiro et al. 2004). Consequently, the Pitx1 proteins encoded by the marine and freshwater populations were the same. However, the location of the gene's expression was drastically different between populations. As in other vertebrates, *Pitx1* was expressed in the developing pelvis of marine larvae. In contrast, expression was greatly reduced or absent in the pelvic region of freshwater stickleback larvae, yet other regions of normal expression, such as the jaws, were not affected (Shapiro et al. 2004; Shapiro et al. 2006b). Therefore, the change in *Pitx1* was predicted to affect a DNA sequence that regulates when and where the gene is expressed. Chan et al. (2010) confirmed this hypothesis by finding DNA deletions near the *Pitx1* gene in several pelvic-reduced populations. When attached to *Pitx1* and injected into embryos from pelvisless sticklebacks, this regulatory region (also known as an enhancer) was capable of restoring pelvic development, thus verifying that the deletion was critical in the evolution and development of pelvic reduction. In contrast to repeated selection on the same low-plate version of *Eda*, Chan et al. detected *different* deletions near *Pitx1* in different populations, suggesting that pelvic reduction in threespine sticklebacks arose repeatedly by independent mutations in different populations.

A likely factor in the repeated involvement of the *Pitx1* regulatory element, as opposed to mutations in the coding sequence of the gene, is pleiotropy; that is, selection on one trait has the potential to affect development of other traits controlled by the same gene. In mice, the pleiotropic effects of *Pitx1* mutations are especially pronounced: complete inactivation of the gene leads not only to hindlimb anomalies, but also jaw and brain deformities (Lanctôt et al. 1999). In contrast, the pelvis-specific *regulatory* mutation in sticklebacks yields an adaptive phenotype that is specific to one trait while leaving other developmental roles of *Pitx1* intact (Shapiro et al. 2004; Chan et al. 2010).

Pelvic reduction is not limited to a single species of stickleback. The ninespine stickleback (*Pungitius pungitius*) diverged from the threespine stickleback at least 10 million years ago, yet these two species have a similar history of postglacial freshwater colonization and repeated evolution of pelvic reduction (Aldenhoven et al. 2010). Based on one study of the ninespine stickleback (Northwest Territories, Canada), *Pitx1* appears to play a role in pelvic reduction in this species as well (Shapiro et al. 2006a). These results in extant, genetically tractable stickleback species might hold clues about mechanisms of pelvic reduction in other species as well. For example, the extensive fossil record of *Gasterosteus doryssus*, an extinct relative of the threespine stickleback, documents the repeated evolution of pelvic reduction in a Miocene population (Bell 1974b; Bell et al. 1985; Bell 1988). As in modern threespine sticklebacks, pelvic reduction in *G. doryssus* shows a pronounced left-side bias, a morphological signature of *Pitx1*-mediated changes (Shapiro et al. 2004; Shapiro et al. 2006a). This morphological trend extends beyond sticklebacks, as pelvic remnants in manatees also show a left-side bias (Shapiro et al. 2006a). The genetic basis of hindlimb reduction in manatees is not

known, but this shared morphological signature of *Pitx1*-mediated reduction provides clues about the molecular mechanisms involved. Together, these examples show that genetics in one species can potentially generate hypotheses for study in other, less genetically tractable species.

Pitx1 probably does not universally play a major role in pelvic reduction, however. In another population of ninespine sticklebacks (Point MacKenzie, Alaska), the major QTL for pelvic reduction is clearly not *Pitx1* (Shapiro et al. 2009). This result suggests that ninespine stickleback populations use both the same and different genetic mechanisms as threespine sticklebacks to converge on the same pelvic phenotype.

Body Shape Variation

Sticklebacks from a variety of habitats exhibit enormous variation in overall body shape. The ancestral marine form is generally large and streamlined with a deep body and head, long fins, and a narrow caudal region. These adaptations are thought to be optimal for navigating open water (Walker 1997; Walker and Bell 2000; Spoljaric and Reimchen 2007; Albert et al. 2008). Freshwater populations, particularly those that inhabit littoral regions and feed on macroinvertebrates, generally have bodies that are short and deep, with shorter fins and a wider caudal region, resulting in a more maneuverable body that is better suited to foraging and evading predators in a complex habitat (Webb 1982; Walker 1997; Walker and Bell 2000; Spoljaric and Reimchen 2007).

While many studies have highlighted recurring trends in body shape and their link to particular habitats, less is known about the genetic architecture of these changes (reviewed in Reid and Peichel, 2010). To address this shortcoming, Albert et al. (2008) used a cross between marine and freshwater fish to conduct QTL mapping for body and

head shape. Perhaps not surprisingly, they found that the genetic architecture of body shape is more complex than discrete traits such as plate variation and pelvic reduction. However, similar to discrete traits, the same genomic regions underlie similar body shape traits in different populations. For example, some of the same chromosome regions influence differences not only between marine and freshwater populations, but also between semi-isolated benthic and limnetic populations that occur within several lakes (Gow et al. 2006; Reid and Peichel 2010).

Collectively, these studies suggest that similar suites of shape changes are key transformations in adaptation to new freshwater habitats, and similar suites of genes might govern these repeated changes species-wide (also see Hohenlohe et al. 2010; Jones et al. 2012a,b).

Summary

Molecular genetic studies of microevolutionary transformations in sticklebacks provide important insights into general trends underlying the molecular basis of a classic adaptive radiation. First, dramatic phenotypic changes such as pelvis and armor reduction can result largely from changes at a few genetic loci (e.g., *Pitx1* and *Eda*, respectively, plus a modest number of loci of small effect). Furthermore, repeated evolution of the same trait can result from repeated selection on a common ancestral chromosome segment (lateral armor evolution and *Eda*) or independent mutations in the same gene (pelvic evolution and *Pitx1*). However, comparisons across stickleback species suggest that these mechanisms are not necessarily universal. Other adaptive changes, such as body shape modifications that characterize populations in different habitats, have a more complex genetic architecture, yet still repeatedly involve a similar suite of genomic regions.

Mexican Cavefish (Family Characidae, *Astyanax mexicanus*)

Introduction

As with freshwater habitat specialization in sticklebacks, cave specialization has resulted in the repeated evolution of similar traits across diverse lineages of metazoans, including teleost fishes. Constructive traits that are common in cave-dwelling animals include increased numbers of taste buds, increased fat storage, larger egg size, and more sensitive nonvisual sensory systems (Culver 1982); regressive traits, such as loss of eyes and pigmentation, have evolved repeatedly across phyla as well.

The Mexican cavefish (*Astyanax mexicanus*) is an ideal model to study the genetic basis of cave phenotypes in vertebrates. Multiple populations within this species have converged on similar phenotypes, providing another opportunity to test whether the same or different genetic mechanisms underlie repeated morphological changes. At least 30 populations of *A. mexicanus* are distributed across northeastern Mexico (Hubbs and Innis 1936; Wilkens and Burns 1972; Mitchell et al. 1977; Espinasa et al. 2001), and phylogenetic analyses suggest that the cave form does not have a single evolutionary origin (Espinasa 2001; Dowling et al. 2002; Strecker et al. 2003; Strecker et al. 2004).

Pigmentation Variation

In the darkness of a cave environment, the usual roles of pigmentation (camouflage, mate selection, etc.) are no longer relevant and the loss of pigmentation has occurred in cave dwelling species across phyla. However, the adaptive significance (if any) of this phenotype in cavefish and other cave animals is still unclear. Pigmentation variation in cavefish encompasses a number of distinct phenotypes, including complete albinism, pigmentation reduction, and decreased melanophore number, each with a

distinct genetic architecture.

Albinism has long been known to be controlled by a single major locus and possibly the same gene in multiple populations (Sadoglu 1957; Sadoglu and McKee 1969; Wilkens 1988). QTL mapping in cavefish led to the discovery of a deletion in the *Oca2* gene that underlies albinism in the Pachón population (Protas et al. 2006) (Figure 3.3, A–C). *Oca2* encodes a key protein in melanin synthesis, and mutations in this gene also cause albinism in both humans and mice (Rinchik et al. 1993; Yi et al. 2003). Albinism in a second cavefish population, Molino, is also due to a deletion in *Oca2*, but this deletion is distinct from the Pachón version and therefore must have arisen independently (Protas et al. 2006). Albinism in a third population, Japonés, probably results from a regulatory mutation in the same gene as no coding changes were identified (Protas et al. 2006). Hence, as with *Pitx1* and pelvic reduction in sticklebacks, different mutations in the same gene led to similar phenotypes in different populations.

Another pigment-reduction phenotype, *brown* (characterized by brown instead of black eyes and reduced melanophore number), results from mutations in the *Melanocortin-1 receptor* (*Mc1r*) gene (Figure 3.3, D–F). *Mc1r* encodes a receptor protein expressed in pigment-producing cells, and its activity can regulate melanin content and melanocyte dispersal in fish (Richardson et al. 2008; Tezuka et al. 2011). Like *Oca2* and albinism, the *brown* phenotype results from more than one mutation in different cavefish populations, although at least one of these mutations has probably spread to several populations (Gross et al. 2009).

Together, these examples of pigment variation illustrate that convergent phenotypes can occur by independent mutations in the same genes (similar to the

repeated evolution of pelvic reduction in sticklebacks) and by selection on standing genetic variants (similar to repeated evolution of armor phenotypes in sticklebacks). In cavefish, independent deletions in the coding region of *Oca2*, as well as a possible regulatory mutation, have all been implicated in albinism. Likewise, independent mutations in *Mc1r* led to repeated evolution of the *brown* phenotype, perhaps by a combination of selection on mutant alleles that originated in the surface population, and *de novo* mutations in different cave populations (Gross et al. 2009).

Eye Loss

One of the most dramatic changes in cavefish relative to their surface-dwelling relatives is severe eye reduction (Figure 3.3, A–C). During embryonic development in cavefish, eyes begin to form but eventually stall and degenerate, beginning with the lens (Cahn 1958; Yamamoto et al. 2004). However, transplanting a surface fish lens into a developing cavefish eye can halt degeneration, demonstrating that this structure is a critical signaling center in eye development (Jeffery and Martasian 1998; Yamamoto and Jeffery 2000; Strickler et al. 2007b).

Genetic and developmental experiments suggest that between 6 and 12 genes contribute to eye regression in cavefish (Wilkens 1988; Protas et al. 2007) and that the same genetic mechanisms do not underlie regression in all cave populations (Wilkens 1971; Wilkens and Strecker 2003; Borowsky 2008). This complex trait probably entails genetic pathways that control cell death and proliferation (Protas et al. 2007; Strickler et al. 2007a; Gross et al. 2008), response to environmental stress (Hooven et al. 2004), photoreceptor development (Kozmik 2008; Strickler 2009), and morphogenesis (Jeffery and Martasian 1998; Yamamoto et al. 2004; Strickler 2009). In summary, eye

degeneration in cavefish is probably not under simple genetic control. Although several specific genes have been shown to affect eye development in this species, no specific mutations have yet been identified that correlate with the eyeless phenotype in any cave population.

Selection, Neutral Mutation, and Pleiotropy

While it is intuitive to envision natural selection driving the acquisition of heightened sensory traits such as increased taste bud number and increased sensitivity to vibrations in a cave environment, the adaptive consequences of eye and pigment loss are less clear. Perhaps unnecessary structures in a dark environment, such as the eye, are a liability; for example, eyes could be targets for predators, injury, or infection (Poulson 1963; Poulson and White 1969; Culver 1982; Jeffery 2005). Alternatively, neutral mutation could explain eye and pigment loss (Kimura and Ohta 1971; Culver 1982; Wilkens 1988). In a dark environment, otherwise deleterious mutations in pigment and eye developmental pathways might not be selected against, as long as they do not result in other disadvantageous phenotypes. Therefore, given sufficient time, pathways involved in eye and pigment development could accumulate enough mutations for the associated structures to be lost. Interestingly, in genetic crossing experiments, cave alleles tend only to contribute to decreases in eye size, consistent with selection on eye regression, while cave alleles contribute to both increases and decreases in number of melanophores, suggesting drift might play a central role in pigmentation traits (Protas et al. 2007).

The loss of eyes and pigmentation in cavefish might also result from pleiotropy. Genetic and experimental evidence suggest that eye reduction might be a secondary effect of selection on alleles that are advantageous in the cave environment for increased

gustatory or mechanical sensitivity (Yamamoto et al. 2004, 2009; Yoshizawa et al. 2010, 2013; Borowsky 2013). For example, in hybrid crosses between cave and surface fish, the number of taste buds is inversely correlated with eye size (Yamamoto et al. 2009). A compelling example of this effect on the developmental level comes from the gene *Sonic hedgehog* (*Shh*), which is expressed in the oral-pharyngeal region and the developing taste buds of both cave and surface forms. When this gene is experimentally overexpressed in both forms, embryos develop wider jaws and more taste buds, as well as smaller eyes (Yamamoto et al. 2004, 2009).

Summary

As in sticklebacks, genetic dissection of derived traits in cavefish demonstrates that dramatic phenotypes can potentially fall under the control of a modest number of genomic regions of large effect. Furthermore, these studies also show that similar phenotypes can arise through independent mutations in the same genes: *Oca2* and *Mc1r* underlie pigmentation variation in several populations, but different populations carry different mutations. Derived pigmentation traits in cavefish can also result from either coding or regulatory mutations: at least one population of albino cavefish probably harbors a regulatory mutation in *Oca2*, while most other albino populations have coding changes that lead to a decrease or loss of function. Other phenotypes, such as eye loss, are genetically more complicated and are probably the result of changes in multiple genes.

Although great strides are being made to identify the genetic basis of derived traits, these data do not necessarily lead directly to an understanding of the adaptive significance of phenotypes. Both pigment and eye reduction might result from positive selection for these traits, neutral mutation, or pleiotropy as the result of selection on other,

as yet unknown, adaptive phenotypes.

Cichlids (Family Cichlidae)

Introduction

Cichlids, a third example of a morphologically diverse and speciose group of teleosts, inhabit lakes throughout Central and South America, Madagascar, India, and Africa. Several lakes throughout this range include classic examples of rapid adaptive radiations. Two especially notable cases occur in the African rift lakes, where more than 500 species in Lake Victoria and over 700 species in Lake Malawi have arisen within the last 1 million years after multiple colonization events and hybridization (Banister and Clarke 1980; Meyer et al. 1990; Owen et al. 1990; Meyer 1993; Kocher et al. 1995; Turner et al. 2001; Joyce et al. 2011). Within a single lake, these species occupy habitats from shallow water to depths of over 100 meters. Different species also have diverse feeding strategies from generalist fish, zooplankton, and algae feeders to specialized crab, snail, and scale eaters (reviewed in Turner 2007). Furthermore, similar strategies have arisen multiple times, providing another opportunity to examine the genetic basis of convergence in adaptively relevant phenotypes (Kocher et al. 1993) (Figure 3.4, a). Like sticklebacks and cavefish, genetic mapping of derived traits in cichlids is greatly facilitated by the ability of many distinct forms to interbreed and produce fertile offspring in a laboratory setting.

Feeding Morphology

Some of the best-studied adaptive traits in cichlids involve craniofacial structures. Different cichlid species have evolved to feed on an enormous variety of food types, and

this diversification has produced a wide range of specialized head, jaw, and tooth morphologies (Albertson and Kocher 2006) (Figure 3.4). Genetic control of jaw and head morphology is highly complex and involves at least 40 chromosome regions, many of them affecting multiple elements of the feeding apparatus (Albertson and Kocher 2001; Albertson et al. 2003a, b).

To reduce this complexity, Albertson et al. (2005) specifically examined functionally relevant aspects of jaw morphology in two divergent species. The first species, *Metriaclima zebra*, feeds on algae, diatoms, and plankton from the water column and has a narrow, forward-directed mouth optimized for suction feeding (Ribbink et al. 1983). In contrast, the jaw of *Labeotropheus fuelleborni* is short and square with a downward orientation that allows it to bite algae from rocks while remaining horizontal (Ribbink et al. 1983). One QTL identified in this study included *Bone Morphogenetic Protein 4* (*Bmp4*), a member of a large gene family that also regulates growth and differentiation during craniofacial development in other vertebrates (Abzhanov et al. 2004; Wu et al. 2004). At early developmental stages, the jaws of the suction-feeder *M. zebra* had much lower *Bmp4* expression than the biting-feeder *L. fuelleborni* (Albertson et al. 2005). Interestingly, when Albertson et al. (2005) overexpressed *Bmp4* in the embryos of zebrafish (suction-feeders, like *M. zebra*), the lower jaw shape shifted to a shape more suited for biting (like *L. fuelleborni*). Therefore, the results of experimental developmental studies in the zebrafish model system were consistent with genetic findings in wild cichlid species.

In another study using the same two species, Roberts et al. (2011) implicated the gene *Patched 1* (*Ptch1*)–a receptor in the hedgehog pathway that contributes to dermal

bone development (Abzhanov et al. 2007)–in morphological differences in the lower jaw. Beyond *M. zebra* and *L. fuelleborni*, additional species-specific alleles of *Ptch1* were found in other cichlids with divergent feeding strategies, suggesting that this gene might affect jaw morphology in multiple lineages (Roberts et al. 2011).

Summary

The search for molecular changes that contribute to adaptive changes among cichlid species has thus far identified a small number of genes that contribute to diversity in feeding morphology, a key feature of this group's radiation. However, the genetic basis of variation in feeding structures is complex, with numerous chromosome regions contributing to differences in morphology. As with body shape variation in sticklebacks and eye reduction in cavefish, feeding morphology in cichlids involves several genomic regions that contribute to variation in multiple structures.

Discussion

Genetic Architecture of Derived Traits

The examples outlined above show that the genetic architecture of some major morphological changes can be relatively simple, with large effects produced by changes in only a few genes or genomic regions. Plate and pelvic reduction in sticklebacks, as well as albinism in cavefish, are largely controlled by single major genes. However, some derived traits have a more complex genetic architecture, including changes in stickleback body shape, variation in cichlid jaw morphology, and reduction of the cavefish eye. These contrasting degrees of complexity might represent different temporal stages of morphological transformations. Theoretical models of adaptation by new mutations (as

opposed to selection on standing genetic variation) suggest that a small number of initial mutations lead to large fitness effects, so early adaptive stages can have a simple genetic architecture; subsequently, "modifier" mutations of smaller effect accumulate over time (Orr 1998, 2002). By this model, several examples of genetically simple changes discussed above might reflect very recent transformations, while a more complex architecture could potentially reflect a longer period of trait evolution or selection on a large number of pre-existing genetic variants.

We also note that, in all three teleost examples, several QTL regions control multiple traits. For instance, in sticklebacks, LG4 appears to be a "hotspot" of variation in body shape, lateral plates, and pelvic phenotypes (Colosimo et al. 2004; Shapiro et al. 2004; Albert et al. 2008; Shapiro et al. 2009). In cavefish, thirteen genomic regions are known to influence multiple traits (Protas et al. 2008); these regions could contain multiple genes that affect a suite of traits beneficial to cave-dwellers, or single genes that have pleiotropic effects. Finally, in cichlids, LG5 influences tooth morphology, female sex determination and pigmentation and also contains genes important for color perception (Carleton and Kocher 2001; Albertson et al. 2003a; Streelman et al. 2003; Kocher 2004; Streelman and Albertson 2006). This trend is by no means limited to loci that underlie diversity in fishes; the genetic clustering of QTL that control ecologically relevant traits could allow rapid evolutionary change through linkage of advantageous alleles in many different organisms (e.g., Garber and Quisenberry 1927; Mather 1950; Sheppard 1953; Murray and Clarke 1973; Joron et al. 2006; Joron et al. 2011).

Coding Versus Regulatory Mutations

Among the teleost examples we discuss above, some of the genetic changes are (or are predicted to be) in noncoding regulatory regions of genes, while others directly affect protein-coding sequences, which in turn can affect protein function. This dichotomy, and relative contributions of each type of mutation to evolutionary change in general, has sparked considerable interest in the recent evolutionary genetics literature (e.g., Hoekstra and Coyne 2007; Wray 2007; Carroll 2008; Stern and Orgogozo 2008). While it is clear that not all evolutionary change results from *cis*-regulatory mutations, a number of hypotheses have been put forth to explain why these noncoding mutations might be a primary driver of evolutionary change, especially morphological change. One compelling argument centers on the modularity of regulatory regions (reviewed in Carroll 2008). Modularity refers to the semi-independent function of each *cis*-regulatory element with respect to other *cis*-regulatory elements. Therefore, a mutation in one of several regulatory regions of a gene can affect gene expression in only a subset of tissues or developmental time points, thereby avoiding potentially detrimental side effects on other developmental processes (pleiotropy). The potential importance of regulatory changes has been appreciated since the description of bacterial operons by Jacob and Monod (1961), and *cis*-regulatory changes are clearly important in morphological, physiological, and behavioral evolution (reviewed in Wray, 2007).

An argument against the dominance of *cis*-regulatory changes in evolutionary change is that there are currently more confirmed examples of coding changes, but this could simply be because coding mutations are much easier to identify than regulatory mutations (reviewed in Stern and Orgogozo, 2008). However, the pace of discovery (or

implication) of *cis*-regulatory changes has recently begun to closely track the discovery of coding changes (Stern and Orgogozo 2008). In summary, both coding and regulatory mutations have the potential to contribute to significant evolutionary transformations, and ongoing work in fishes and other organisms will further elucidate general trends, if any exist.

Convergent Evolution

Teleosts exhibit repeated evolution of similar phenotypes among different populations within a species, and in some cases, between species. In many populations of threespine sticklebacks, lateral armor reduction evolved by repeated selection on a standing variant of the *Eda* locus. In contrast, other convergent evolutionary changes are the products of different mutations in the same genes. For example, different mutations in *Pitx1* underlie pelvic reduction in several populations of threespine sticklebacks, and *Oca2* and *Mc1r* mutations differ among cavefish populations with similar pigmentation phenotypes.

Comparisons *between* stickleback species also yield novel insights about convergent phenotypes. For example, pelvic reduction in at least one population of ninespine sticklebacks probably results from changes to *Pitx1*, just as in threespine sticklebacks (Shapiro et al. 2006a). However, in another population of ninespine sticklebacks, pelvic reduction is controlled by a genomic region distinct from *Pitx1*; QTL for other skeletal traits (including lateral armor) and sex determination also differ between the two species (Shapiro et al. 2009). Therefore, a multispecies approach can be particularly informative in dissecting a broad range of genetic mechanisms underlying similar phenotypes.

Future Directions

Biologists are intensely interested in how vertebrates undergo transformations both great and small, yet we know remarkably little about the genetic basis of phenotypic change. In several examples above, QTL results were leveraged to fine-map and functionally test specific candidate genes for the evolution of derived traits. While these cases are exciting, it is important to note that they are also currently the exceptions– mapping traits to the gene level and demonstrating functional consequences of mutations is still uncommon.

Traits with a simple genetic architecture are easier to analyze than those with more genetic complexity, and many traits that have been examined in natural populations of teleosts and other organisms are ones that are relatively easy to see and quantify. Therefore, observable and relatively simple traits are preferentially studied, and we have a poorer understanding of complex anatomical, physiological, and behavioral traits that are undoubtedly important for evolutionary transformations (Rockman 2012).

New genomic tools, and the ability to compare dozens of genomes simultaneously, can help identify signatures of selection in suites of genes that affect traits that are not easily visualized. Recent studies, perhaps most notably in sticklebacks (Hohenlohe et al. 2010; Jones et al. 2012a; Jones et al. 2012b), have taken this "bottom-up" approach to identify genomic regions under selection in marine versus freshwater environments, as well as in benthic versus limnetic freshwater habitats. With precipitous drops in the cost of DNA sequencing and generation of genetic resources, we expect that techniques pioneered for a limited number of species will become widely available to investigate important evolutionary transformations in other vertebrates as well.

Glossary

Allele: Variants of a given gene or marker.

Coding mutation: A change in DNA sequence that occurs in a part of a gene that codes for a protein.

Genetic architecture: A general description of how traits are controlled by genotypes. For example, genetic architecture includes the number and location of genes that underlie a trait, as well as the number of alleles at these loci and the interactions among them.

Genetic marker: A DNA sequence that shows variability among individuals, and thus the inheritance of different alleles can be traced from one generation to the next. Examples include single nucleotide polymorphisms (SNPs) and microsatellites.

Genotype: The genetic make-up of an organism.

Linkage group: A group of genes or genetic markers that reside on the same chromosome. Genes or markers that are physically close to one another tend to be inherited together; as a result, markers can be ordered by tracking transmission from one generation to the next (also called genetic mapping). The sum of linkage groups comprises a linkage map.

Locus (plural: loci): The location of a gene or DNA sequence on a chromosome or linkage group.

Phenotype: The observable characteristics of an organism.

Pleiotropy: When one gene affects more than one trait or developmental process.

QTL (quantitative trait locus): A genomic region that contributes to variation in a trait. Quantitative traits are typically controlled by multiple loci.

QTL mapping: An experimental approach that often begins by crossing strains of

organisms that differ in a trait or traits of interest. Molecular markers across the genome are used to track the co-inheritance of genotypes and phenotypes of offspring. Correlations between the trait(s) of interest and molecular markers are assessed (see Figure 2, and (Miles and Wayne 2008).

Regulatory (*cis*-) mutation: A change in DNA sequence that affects a region controlling the level or location of expression of a gene, but (typically) does not affect the protein encoded by the gene (see also Wray 2007, Carroll 2008).

References

Abzhanov, A., M. Protas, B. R. Grant, P. R. Grant, and C. J. Tabin. 2004. Bmp4 and morphological variation of beaks in Darwin's finches. Science 305:1462–1465.

Abzhanov, A., S. J. Rodda, A. P. McMahon, and C. J. Tabin. 2007. Regulation of skeletogenic differentiation in cranial dermal bone. Development 134:3133–3144.

Albert, A. Y., S. Sawaya, T. H. Vines, A. K. Knecht, C. T. Miller, B. R. Summers, S. Balabhadra, D. M. Kingsley, and D. Schluter. 2008. The genetics of adaptive shape shift in stickleback: pleiotropy and effect size. Evolution 62:76–85.

Albertson, R. C. and T. D. Kocher. 2001. Assessing morphological differences in an adaptive trait: a landmark-based morphometric approach. J. Exp. Zool. 289:385–403.

Albertson, R. C. and T. D. Kocher. 2006. Genetic and developmental basis of cichlid trophic diversity. Heredity (Edinb) 97:211–221.

Albertson, R. C., J. T. Streelman, and T. D. Kocher. 2003a. Directional selection has shaped the oral jaws of Lake Malawi cichlid fishes. Proc. Natl. Acad. Sci. USA 100:5252–5257.

Albertson, R. C., J. T. Streelman, and T. D. Kocher. 2003b. Genetic basis of adaptive shape differences in the cichlid head. J. Hered. 94:291–301.

Albertson, R. C., J. T. Streelman, T. D. Kocher, and P. C. Yelick. 2005. Integration and evolution of the cichlid mandible: the molecular basis of alternate feeding strategies. Proc. Natl. Acad. Sci. USA 102:16287–16292.

Aldenhoven, J. T., M. A. Miller, P. S. Corneli, and M. D. Shapiro. 2010. Phylogeography

of ninespine sticklebacks (Pungitius pungitius) in North America: glacial refugia and the origins of adaptive traits. Mol. Ecol. 19:4061–4076.

Avise, J. C. 1976. Genetics of plate morphology in an unusual poulation of threespine sticklebacks (*Gasterostus aculeatus*). Genet. Res. 27:33–46.

Banbura, J. 1994. A new model of lateral plate morph inheritance in the threespine stickleback, *Gasterosteus aculeatus*. Theor. Appl. Genet. 88:871–876.

Banister, K. E. and M. A. Clarke. 1980. A revision of the large Barbus (Pisces, Cyprinidae) of Lake Malawi with a reconstruction of the history of the southern African Rift Valley lakes. J. Nat. Hist. 14:483.

Barrett, R. D. 2010. Adaptive evolution of lateral plates in three-spined stickleback Gasterosteus aculeatus: a case study in functional analysis of natural variation. J. Fish. Biol. 77:311–328.

Bell, A. M., G. Orti, J. A. Walker, and J. P. Koenings. 1993. Evolution of pelvic reduction in threespine stickleback fish-a test of competing hypotheses. Evolution 47:906–914.

Bell, M. A. 1974a. Reduction and loss of the pelvic girdle in Gasterosteus (Pisces): a case of parallel evolution. Nat. Hist. Mus. L.A. Contrib. Sci. 257:1–36.

Bell, M. A. 1974b. Reduction and loss of the pelvic girdle in *Gasterosteus* (Pisces): a case of parallel evolution. Natural History Museum of Los Angeles County Contributions in Science 257:1–36.

Bell, M. A. 1987. Interacting evolutionary constrains in pelvic reduction of threespine sticklebacks, *Gasterosteus aculeatus* (Pisces, Gasterosteidae). Biol. J. Linn. Soc. 31:347–382.

Bell, M. A. 1988. Stickleback fishes: bridging the gap between population biology and paleobiology. Trends Ecol. Evol. 3:320–325.

Bell, M. A., W. E. Aguirre, and N. J. Buck. 2004. Twelve years of contemporary armor evolution in a threespine stickleback population. Evolution 58:814–824.

Bell, M. A., J. V. Baumgartner, and E. C. Olson. 1985. Patterns of temporal change in single morphological characters of a Miocene stickleback fish. Paleobiology 11:258–271.

Bell, M. A. and S. A. Foster. 1994. The Evolutionary Biology of the Threespine Stickleback. Oxford Univ. Press, Oxford.

Bell, M. A. and G. Orti. 1994. Pelvic reduction in threespine stickleback from Cook Inlet

lakes: geographic distribution and intrapopulation variation. Copeia 1994:314–325.

Bernatchez, L. and C. C. Wilson. 1998. Comparative phylogeography of Nearctic and Palearctic fishes. Mol. Ecol. 7:431–452.

Blouw, D. M. and G. J. Boyd. 1992. Inheritance of reduction, loss, and asymmetry of the pelvis of *Pungitius pungitius* (ninespine stickleback). Heredity 68:33–42.

Borowsky, R. 2008. Restoring sight in blind cavefish. Curr. Biol. 18:R23–24.

Borowsky, R. 2013. Eye regression in blind Astyanax cavefish may facilitate the evolution of an adaptive behavior and its sensory receptors. BMC Biol. 11:81.

Bourgeois, J. F., D. M. Blouw, and M. A. Bell. 1994. Multivariate analysis of geographic covariance between phenotypes and environments in the threespine stickleback, *Gasterosteus aculeatus*. Can. J. Zool. 72:1497–1509.

Cahn, P. H. 1958. Comparative optic development in *Astyanax mexicanus* and two of its blind cave derivatives. Bull. Am. Mus. Nat. Hist. 115:75–112.

Campbell, R. N. and R. B. Williamson. 1979. The fishes of inland waters in the Outer Hebrides. Proc. R. Soc. Edinb. 77B:377–393.

Carleton, K. L. and T. D. Kocher. 2001. Cone opsin genes of african cichlid fishes: tuning spectral sensitivity by differential gene expression. Mol. Biol. Evol. 18:1540–1550.

Carroll, S. B. 2008. Evo-devo and an expanding evolutionary synthesis: a genetic theory of morphological evolution. Cell 134:25–36.

Chan, Y. F., M. E. Marks, F. C. Jones, G. Villarreal, Jr., M. D. Shapiro, S. D. Brady, A. M. Southwick, D. M. Absher, J. Grimwood, J. Schmutz, R. M. Myers, D. Petrov, B. Jonsson, D. Schluter, M. A. Bell, and D. M. Kingsley. 2010. Adaptive evolution of pelvic reduction in sticklebacks by recurrent deletion of a Pitx1 enhancer. Science 327:302–305.

Coad, B. W. 1983. Plate morphs in freshwater samples of Gasterosteus aculeatus from Arctic and ATlantic Canada: complementary comments on a recent contribution. Can. J. Zool. 61:1174–1177.

Colosimo, P. F., K. E. Hosemann, S. Balabhadra, G. Villarreal, Jr., M. Dickson, J. Grimwood, J. Schmutz, R. M. Myers, D. Schluter, and D. M. Kingsley. 2005. Widespread parallel evolution in sticklebacks by repeated fixation of Ectodysplasin alleles. Science 307:1928–1933.

Colosimo, P. F., C. L. Peichel, K. Nereng, B. K. Blackman, M. D. Shapiro, D. Schluter, and D. M. Kingsley. 2004. The genetic architecture of parallel armor plate reduction in threespine sticklebacks. PLos Biol. 2:E109.

Cresko, W. A., A. Amores, C. Wilson, J. Murphy, M. Currey, P. Phillips, M. A. Bell, C. B. Kimmel, and J. H. Postlethwait. 2004. Parallel genetic basis for repeated evolution of armor loss in Alaskan threespine stickleback populations. Proc. Natl. Acad. Sci. USA 101:6050–6055.

Culver, D. C. 1982. Cave life : evolution and ecology. Harvard University Press, Cambridge, MA.

Darwin, C. 1859. On the Origin of Species by Means of Natural Selection. John Murray, London.

Dowling, T. E., D. P. Martasian, and W. R. Jeffery. 2002. Evidence for multiple genetic forms with similar eyeless phenotypes in the blind cavefish, Astyanax mexicanus. Mol. Biol. Evol. 19:446–455.

Edge, T. A. and B. W. Coad. 1983. Reduction of the pelvic skeleton in the three-spined stickleback *Gasterosteus aculeatus* in 2 lakes of Quebec Canada. Can. Field. Nat. 97:334–336.

Espinasa, L., Borowsky, R.B. 2001. Origin and relationships of cave populations of the blind Mexican tetra *Astyanax fasciatus*, in the Sierra de El Abra. Environ. Biol. Fishes 62:233–237.

Espinasa, L., P. Rivas-Manzano, and H. Espinosa Pérez. 2001. A new blind cave fish population of the genus *Astyanax*: geography, morphology and behavior. Environ. Biol. Fishes 62:329–344.

Garber, R. J. and K. S. Quisenberry. 1927. The inheritance of length of syle in buckwheat. J. Agric. Res. 34:181–183.

Giles, N. 1983. The possible role of environmental calcium levels during the evolution of phenotypic diversity in Outer-Hebridean populations of the three-spined stickleback, *Gasterosteus aculeatus*. J. Zool. 199:535.

Gow, J. L., C. L. Peichel, and E. B. Taylor. 2006. Contrasting hybridization rates between sympatric three-spined sticklebacks highlight the fragility of reproductive barriers between evolutionarily young species. Mol. Ecol. 15:739–752.

Gross, H. P. 1978. Natural selection by predators on the defensive apparatus of the three-spined stickleback, *Gasterosteus aculeatus* L. Can. J. Zool. 56:398–413.

Gross, J. B., R. Borowsky, and C. J. Tabin. 2009. A novel role for Mc1r in the parallel evolution of depigmentation in independent populations of the cavefish Astyanax mexicanus. PLoS Genet. 5:e1000326.

Gross, J. B., M. Protas, M. Conrad, P. E. Scheid, O. Vidal, W. R. Jeffery, R. Borowsky, and C. J. Tabin. 2008. Synteny and candidate gene prediction using an anchored linkage map of Astyanax mexicanus. Proc. Natl. Acad. Sci. USA 105:20106–20111.

Gurnett, C. A., F. Alaee, L. M. Kruse, D. M. Desruisseau, J. T. Hecht, C. A. Wise, A. M. Bowcock, and M. B. Dobbs. 2008. Asymmetric lower-limb malformations in individuals with homeobox PITX1 gene mutation. Am. J. Hum. Genet. 83:616–622.

Hagen, D. W. and L. G. Gilbertson. 1972. Geographic variation and environmental selection in *Gasterosteus aculeaus* L. in the Pacific northwest, America. Evolution 26:32–51.

Hagen, D. W. and L. G. Gilbertson. 1973a. The genetics of plate morphs in freshwater threespine sticklebacks. Heredity 31:75–84.

Hagen, D. W. and L. G. Gilbertson. 1973b. Selective predation and the intensity of selection acting upon the lateral plates of threespine stickleacks. Heredity 30:75–84.

Hagen, D. W. and G. E. E. Moodie. 1982. Polymorphism for plate morphs in *Gasterosteus aculeatus* on the east coast of Canada and an hypothesis for their global distribution. Can. J. Zool. 60:1032–1042.

Harris, M. P., N. Rohner, H. Schwarz, S. Perathoner, P. Konstantinidis, and C. Nusslein-Volhard. 2008. Zebrafish eda and edar mutants reveal conserved and ancestral roles of ectodysplasin signaling in vertebrates. PLoS Genet. 4:e1000206.

Heuts, M. C. 1947. The phenotypical variability of *Gasterosteus aculeatus* (L.) populations in Belgium; its bearing on the general geographical variability of the species. Standaard-Boekhandel, Antwerpen.

Hewitt, G. 2000. The genetic legacy of the Quaternary ice ages. Nature 405:907–913.

Hoekstra, H. E. and J. A. Coyne. 2007. The locus of evolution: evo devo and the genetics of adaptation. Evolution 61:995–1016.

Hohenlohe, P. A., S. Bassham, P. D. Etter, N. Stiffler, E. A. Johnson, and W. A. Cresko. 2010. Population genomics of parallel adaptation in threespine stickleback using sequenced RAD tags. PLoS Genet. 6:e1000862.

Hoogland, R. D., D. Morris, and N. Tinbergen. 1957. The spines of sticklebacks (*Gasterosteus* and *Pygosteus*) as means of defense against predators (*Perca* and *Esox*). Behaviour 10:205–230.

Hooven, T. A., Y. Yamamoto, and W. R. Jeffery. 2004. Blind cavefish and heat shock protein chaperones: a novel role for hsp90alpha in lens apoptosis. Int. J. Dev. Biol. 48:731–738.

Hubbs, C. L. and W. T. Innis. 1936. The first known blind fish of the family Characidae: a new genus from Mexico. Occas. Papers Mus. Zool. Univ. Michigan 342:1–7.

Huxley, T. H. H. 1860. The Origin of Species. Westminster Review 17:541–570.

Jacob, F. and J. Monod. 1961. Genetic regulatory mechanisms in the synthesis of proteins. J. Mol. Biol. 3:318–356.

Jeffery, W. R. 2005. Adaptive evolution of eye degeneration in the Mexican blind cavefish. J. of Hered. 96:185–196.

Jeffery, W. R. and D. P. Martasian. 1998. Evolution of eye regression in the cavefish *Astyanax*: apoptosis and the *Pax-6* gene. Amer. Zool. 38:685–696.

Jones, F. C., Y. F. Chan, J. Schmutz, J. Grimwood, S. D. Brady, A. M. Southwick, D. M. Absher, R. M. Myers, T. E. Reimchen, B. E. Deagle, D. Schluter, and D. M. Kingsley. 2012a. A genome-wide SNP genotyping array reveals patterns of global and repeated species-pair divergence in sticklebacks. Curr. Biol. 22:83–90.

Jones, F. C., M. G. Grabherr, Y. F. Chan, P. Russell, E. Mauceli, J. Johnson, R. Swofford, M. Pirun, M. C. Zody, S. White, E. Birney, S. Searle, J. Schmutz, J. Grimwood, M. C. Dickson, R. M. Myers, C. T. Miller, B. R. Summers, A. K. Knecht, S. D. Brady, H. Zhang, A. A. Pollen, T. Howes, C. Amemiya, J. Baldwin, T. Bloom, D. B. Jaffe, R. Nicol, J. Wilkinson, E. S. Lander, F. Di Palma, K. Lindblad-Toh, and D. M. Kingsley. 2012b. The genomic basis of adaptive evolution in threespine sticklebacks. Nature 484:55–61.

Joron, M., L. Frezal, R. T. Jones, N. L. Chamberlain, S. F. Lee, C. R. Haag, A. Whibley, M. Becuwe, S. W. Baxter, L. Ferguson, P. A. Wilkinson, C. Salazar, C. Davidson, R. Clark, M. A. Quail, H. Beasley, R. Glithero, C. Lloyd, S. Sims, M. C. Jones, J. Rogers, C. D. Jiggins, and R. H. ffrench-Constant. 2011. Chromosomal rearrangements maintain a polymorphic supergene controlling butterfly mimicry. Nature 477:203–206.

Joron, M., R. Papa, M. Beltran, N. Chamberlain, J. Mavarez, S. Baxter, M. Abanto, E. Bermingham, S. J. Humphray, J. Rogers, H. Beasley, K. Barlow, R. H. ffrench-Constant, J. Mallet, W. O. McMillan, and C. D. Jiggins. 2006. A conserved supergene locus controls colour pattern diversity in Heliconius butterflies. PLos

Biol. 4:e303.

Joyce, D. A., D. H. Lunt, M. J. Genner, G. F. Turner, R. Bills, and O. Seehausen. 2011. Repeated colonization and hybridization in Lake Malawi cichlids. Curr. Biol. 21:R108–109.

Kangas, A. T., A. R. Evans, I. Thesleff, and J. Jernvall. 2004. Nonindependence of mammalian dental characters. Nature 432:211–214.

Kimura, M. and T. Ohta. 1971. Theoretical aspects of population genetics. Princeton University Press, Princeton, N.J.

Kitano, J., D. I. Bolnick, D. A. Beauchamp, M. M. Mazur, S. Mori, T. Nakano, and C. L. Peichel. 2008. Reverse evolution of armor plates in the threespine stickleback. Curr. Biol. 18:769–774.

Kocher, T. D. 2004. Adaptive evolution and explosive speciation: the cichlid fish model. Nat. Rev. Genet. 5:288–298.

Kocher, T. D., J. A. Conroy, K. R. McKaye, and J. R. Stauffer. 1993. Similar morphologies of cichlid fish in Lakes Tanganyika and Malawi are due to convergence. Mol. Phylogenet. Evol. 2:158–165.

Kocher, T. D., J. A. Conroy, K. R. McKaye, J. R. Stauffer, and S. F. Lockwood. 1995. Evolution of NADH dehydrogenase subunit 2 in east African cichlid fish. Mol. Phylogenet. Evol. 4:420–432.

Kozmik, Z. 2008. The role of Pax genes in eye evolution. Brain. Res. Bull. 75:335-339.

Lanctôt, C., A. Moreau, M. Chamberland, M. L. Tremblay, and J. Drouin. 1999. Hindlimb patterning and mandible development require the Ptx1 gene. Development 126:1805–1810.

Lescak, E. A. and F. A. von Hippel. 2011. Selective predation of threespine stickleback by rainbow trout. Ecol. Freshw. Fish 20:308–314.

Marchinko, K. B. 2009. Predation's role in repeated phenotypic and genetic divergence of armor in threespine stickleback. Evolution 63:127–138.

Marcil, A., E. Dumontier, M. Chamberland, S. A. Camper, and J. Drouin. 2003. Pitx1 and Pitx2 are required for development of hindlimb buds. Development 130:45–55.

Mather, K. 1950. The genetical architecture of heterostyly in Primula sinensis. Evolution 4:340–352.

Meyer, A. 1993. Phylogenetic relationships and evolutionary processes in East African cichlid fishes. Trends. Ecol. Evol. 8:279–284.

Meyer, A., T. D. Kocher, P. Basasibwaki, and A. C. Wilson. 1990. Monophyletic origin of Lake Victoria cichlid fishes suggested by mitochondrial DNA sequences. Nature 347:550–553.

Miles, C. M. and M. Wayne. 2008. Quantitative Trait Locus (QTL) Analysis. Nat. Edu. 1:1–8.

Mitchell, R. W., W. H. Russell, and W. R. Elliot. 1977. Mexican eyeless characin fishes, genus *Astyanax*: environment, distribution and evolution. Spec. Publ. Mus. Texas Tech Univ. 12:1–89.

Moodie, G. E. E. 1972. Predation, natural selection and adaptation in an unusual threespine stickleback. Heredity 28:155–167.

Moodie, G. E. E., J. D. McPhail, and D. W. Hagen. 1973. Experimental demonstration of selective predation on *Gasterosteus aculeatus*. Behaviour 47:95–105.

Moodie, G. E. E., and Reimchen, T. 1976. Phenetic variation and habitat differences in *Gasterosteus* populations of the Queen Charlotte Islands. Syst. Zool. 25:49–61.

Murray, J. and B. Clarke. 1973. Supergenes in polymorphic land snails-examples from genus *Partula*. Genetics 74:S188–S189.

Nelson, J. S. and F. M. Atton. 1971. Georgraphic and morphological variation in the presence and absence of the pelvic skeleton in the brook stickleback, Culaea inconstans (Kirtland), in Alberta and Saskatchewan. Can. J. Zool. 49:343–352.

Orr, H. A. 1998. The population genetics of adaptation: the distribution of factors fixed during adaptive evolution. Evolution 52:935–949.

Orr, H. A. 2002. The population genetics of adaptation: the adaptation of DNA sequences. Evolution 56:1317-1330.

Owen, R. B., R. Crossley, T. C. Johnson, D. Tweddle, I. Kornfield, S. Davison, D. H. l. Eccles, and D. E. Engstrom. 1990. Major low levels of Lake Malawi and their implications for speciation rates in cichlid fishes. Proc. R. Soc. London B. 240:519–553.

Poulson, T. L. 1963. Cave adaptation in amblyopsid fishes. Am. Mid. Nat. 70:257–290.

Poulson, T. L. and W. B. White. 1969. The cave environment. Science 165:971–981.

Protas, M., M. Conrad, J. B. Gross, C. Tabin, and R. Borowsky. 2007. Regressive

evolution in the Mexican cave tetra, *Astyanax mexicanus*. Curr. Biol. 17:452–454.

Protas, M., I. Tabansky, M. Conrad, J. B. Gross, O. Vidal, C. J. Tabin, and R. Borowsky. 2008. Multi-trait evolution in a cave fish, Astyanax mexicanus. Evol. Dev. 10:196–209.

Protas, M. E., C. Hersey, D. Kochanek, Y. Zhou, H. Wilkens, W. R. Jeffery, L. I. Zon, R. Borowsky, and C. J. Tabin. 2006. Genetic analysis of cavefish reveals molecular convergence in the evolution of albinism. Nat. Genet. 38:107–111.

Reid, D. T. and C. L. Peichel. 2010. Perspectives on the genetic architecture of divergence in body shape in sticklebacks. Integr. Comp. Biol. 50:1057–1066.

Reimchen, T. E. 1980. Spine deficiency and polymorphism in a population of *Gasterosteus aculeatus*-an adaptation to predators. Can. J. Zool. 58:1232–1244.

Reimchen, T. E. 1983. Structural relationships between spines and lateral plates in threespine stickleback (*Gasterosteus aculeatus*). Evolution 37:931–946.

Reimchen, T. E. 1992. Injuries on stickleback from attacks by a toothed predator (*Oncorhynchus*) and implications for the evolution of lateral plates. Evolution 46:1224.

Reimchen, T. E. 1995. Predator-induced cyclical changes in lateral plate frequencies of *Gasterosteus*. Behaviour 132:1079.

Reist, J. D. 1980. Predation upon pelvic phenotypes of brook stickleback, *Culaea inconstans*, by selected invertebrates. Can. J. Zool. 58:1253–1258.

Ribbink, A. J., A. C. Marsh, C. C. Ribbink, and B. J. Sharp. 1983. A preliminary survey of the cichlid fishes of rocky habitats in Lake Malawi. S. Afr. J. Zool. 18:149–310.

Richardson, J., P. R. Lundegaard, N. L. Reynolds, J. R. Dorin, D. J. Porteous, I. J. Jackson, and E. E. Patton. 2008. mc1r Pathway regulation of zebrafish melanosome dispersion. Zebrafish 5:289–295.

Rinchik, E. M., S. J. Bultman, B. Horsthemke, S. T. Lee, K. M. Strunk, R. A. Spritz, K. M. Avidano, M. T. Jong, and R. D. Nicholls. 1993. A gene for the mouse pink-eyed dilution locus and for human type II oculocutaneous albinism. Nature 361:72–76.

Roberts, R. B., Y. Hu, R. C. Albertson, and T. D. Kocher. 2011. Craniofacial divergence and ongoing adaptation via the hedgehog pathway. Proc. Natl. Acad. Sci. USA 108:13194–13199.

Rockman, M. V. 2012. The QTN program and the alleles that matter for evolution: all

that's gold does not glitter. Evolution 66:1–17.

Sadoglu, P. 1957. A Mendelian gene for albinism in natural cave fish. Experientia 13:394.

Sadoglu, P. and A. McKee. 1969. A second gene that affects eye and body color in Mexican blind cavefish. J. Hered. 60:10-14.

Santini, F., L. J. Harmon, G. Carnevale, and M. E. Alfaro. 2009. Did genome duplication drive the origin of teleosts? A comparative study of diversification in ray-finned fishes. BMC Evol. Biol. 9:194–209.

Schluter, D., E. A. Clifford, M. Nemethy, and J. S. McKinnon. 2004. Parallel evolution and inheritance of quantitative traits. Am. Nat. 163:809–822.

Shapiro, M. D., M. A. Bell, and D. M. Kingsley. 2006a. Parallel genetic origins of pelvic reduction in vertebrates. Proc. Natl. Acad. Sci. USA 103:13753–13758.

Shapiro, M. D., M. E. Marks, C. L. Peichel, B. K. Blackman, K. S. Nereng, B. Jonsson, D. Schluter, and D. M. Kingsley. 2004. Genetic and developmental basis of evolutionary pelvic reduction in threespine sticklebacks. Nature 428:717–723.

Shapiro, M. D., M. E. Marks, C. L. Peichel, B. K. Blackman, K. S. Nereng, B. Jónsson, D. Schluter, and D. M. Kingsley. 2006b. Corrigendum: Genetic and developmental basis of evolutionary pelvic reduction in threespine sticklebacks. Nature 439.

Shapiro, M. D., B. R. Summers, S. Balabhadra, J. T. Aldenhoven, A. L. Miller, C. B. Cunningham, M. A. Bell, and D. M. Kingsley. 2009. The genetic architecture of skeletal convergence and sex determination in ninespine sticklebacks. Curr. Biol. 19:1140–1145.

Sheppard, P. M. 1953. Polymorphism, linkage and the blood groups. Am. Nat. 87:283–294.

Spoljaric, M. A. and T. E. Reimchen. 2007. 10,000 years later: evolution of body shape in Haida Gwaii three-spined stickleback. J. Fish. Biol. 70:1484–1503.

Stern, D. L. and V. Orgogozo. 2008. The loci of evolution: how predictable is genetic evolution? Evolution 62:2155-2177.

Strecker, U., L. Bernatchez, and H. Wilkens. 2003. Genetic divergence between cave and surface populations of Astyanax in Mexico (Characidae, Teleostei). Mol. Ecol. 12:699–710.

Strecker, U., V. H. Faundez, and H. Wilkens. 2004. Phylogeography of surface and cave *Astyanax* (Teleostei) from Central and North America based on cytochrome b sequence data. Mol. Phylogenet. Evol. 33:469–481.

Streelman, J. T. and R. C. Albertson. 2006. Evolution of novelty in the cichlid dentition. J. Exp. Zool. B. Mol. Dev. Evol. 306:216–226.

Streelman, J. T., R. C. Albertson, and T. D. Kocher. 2003. Genome mapping of the orange blotch colour pattern in cichlid fishes. Mol. Ecol. 12:2465–2471.

Strickler, A. G., M. S. Byerly, and W. R. Jeffery. 2007a. Lens gene expression analysis reveals downregulation of the anti-apoptotic chaperone αA-crystallin during cavefish eye degeneration. Dev. Genes. Evol. 217:771–782.

Strickler, A. G., Jeffery, W.R. 2009. Differentially expressed genes identified by cross-species microarray in the blind cavefish *Astyanax*. Int. Zool. 1:99–109.

Strickler, A. G., Y. Yamamoto, and W. R. Jeffery. 2007b. The lens controls cell survival in the retina: Evidence from the blind cavefish *Astyanax*. Dev. Biol. 311:512–523.

Tezuka, A., H. Yamamoto, J. Yokoyama, C. van Oosterhout, and M. Kawata. 2011. The MC1R gene in the guppy (*Poecilia reticulata*): Genotypic and phenotypic polymorphisms. BMC Res. Notes. 4:4–31.

Thesleff, I. and M. L. Mikkola. 2002. Death receptor signaling giving life to ectodermal organs. Sci. STKE 131:pe22.

Turner, G. F. 2007. Adaptive radiation of cichlid fish. Curr. Biol. 17:R827-831.

Turner, G. F., O. Seehausen, M. E. Knight, C. J. Allender, and R. L. Robinson. 2001. How many species of cichlid fishes are there in African lakes? Mol. Ecol. 10:793-806.

Walker, J. A. 1997. Ecological morphology of lacustrine three-spine stickleback *Gasterosteus aculeatus* L. (Gasterosteidae) body shape. Biol. J. Linn. Soc. 61:3–50.

Walker, J. A. and M. A. Bell. 2000. Net evolutionary trajectories of body shape evolution within a microgeographic radiation of threespine sticklebacks. J. Zool. Lond 252:293–302.

Webb, P. W. 1982. Locomotor patters in the evolution of actinopterygian fishes. Am. Zool. 22:329–342.

Wilkens, H. 1971. Genetic interpretation of regressive evolutionary processes: studies of hybrid eyes of two *Astyanax* cave populations (Characidae, Pisces). Evolution 25:530–544.

Wilkens, H. 1988. Evolution and genetics of epigean and cave Astyanax fasciatus (Characidae, Pisces). Evol. Biol. 23:271–367

Wilkens, H. and R. L. Burns. 1972. A new *Anoptichthys* cave population (Characidae, Pisces). Ann. Spéléol. 27:263–270.

Wilkens, H. and U. Strecker. 2003. Convergent evolution of the cavefish *Astyanax* (Charcidae, Telesotei): genetic evidence from reduced eye-size and pigmentation. Biol. J. Linn. Soc. 80:545–554.

Wootton, R. J. 1976. The Biology of the Sticklebacks. Academic, London.

Wray, G. A. 2007. The evolutionary significance of cis-regulatory mutations. Nat. Rev. Genet. 8:206–216.

Wu, P., T. X. Jiang, S. Suksaweang, R. B. Widelitz, and C. M. Chuong. 2004. Molecular shaping of the beak. Science 305:1465–1466.

Yamamoto, Y., M. S. Byerly, W. R. Jackman, and W. R. Jeffery. 2009. Pleiotropic functions of embryonic sonic hedgehog expression link jaw and taste bud amplification with eye loss during cavefish evolution. Dev. Biol. 330:200–211.

Yamamoto, Y. and W. R. Jeffery. 2000. Central role for the lens in cavefish eye degeneration. Science 289:631–633.

Yamamoto, Y., D. W. Stock, and W. R. Jeffery. 2004. Hedgehog signalling controls eye degeneration in blind cavefish. Nature 431:844–847.

Yi, Z., N. Garrison, O. Cohen-Barak, T. M. Karafet, R. A. King, R. P. Erickson, M. F. Hammer, and M. H. Brilliant. 2003. A 122.5-kilobase deletion of the P gene underlies the high prevalence of oculocutaneous albinism type 2 in the Navajo population. Am. J. Hum. Genet. 72:62–72.

Yoshizawa, M., S. Goricki, D. Soares, and W. R. Jeffery. 2010. Evolution of a behavioral shift mediated by superficial neuromasts helps cavefish find food in darkness. Curr. Biol. 20:1631–1636.

Yoshizawa, M., K. E. O'Quin, and W. R. Jeffery. 2013. Evolution of an adaptive behavior and its sensory receptors promotes eye regression in blind cavefish: response to Borowsky (2013). BMC Biol. 11:82–84.

Ziuganov, V. V. 1983. Genetics of osteal plate polymorphism and microevolution of threespine stickleback (*Gasterosteus aculeatus* L). Theor. Appl. Genet. 65:239–246.

Ziuganov, V. V. and A. A. Zotin. 1995. Pelvic girdle polymorphism and reproductive barriers in the ninespine stickleback *Pungitius pungitius* (L.) from northwest Russia. Behaviour 132:1095–1105.

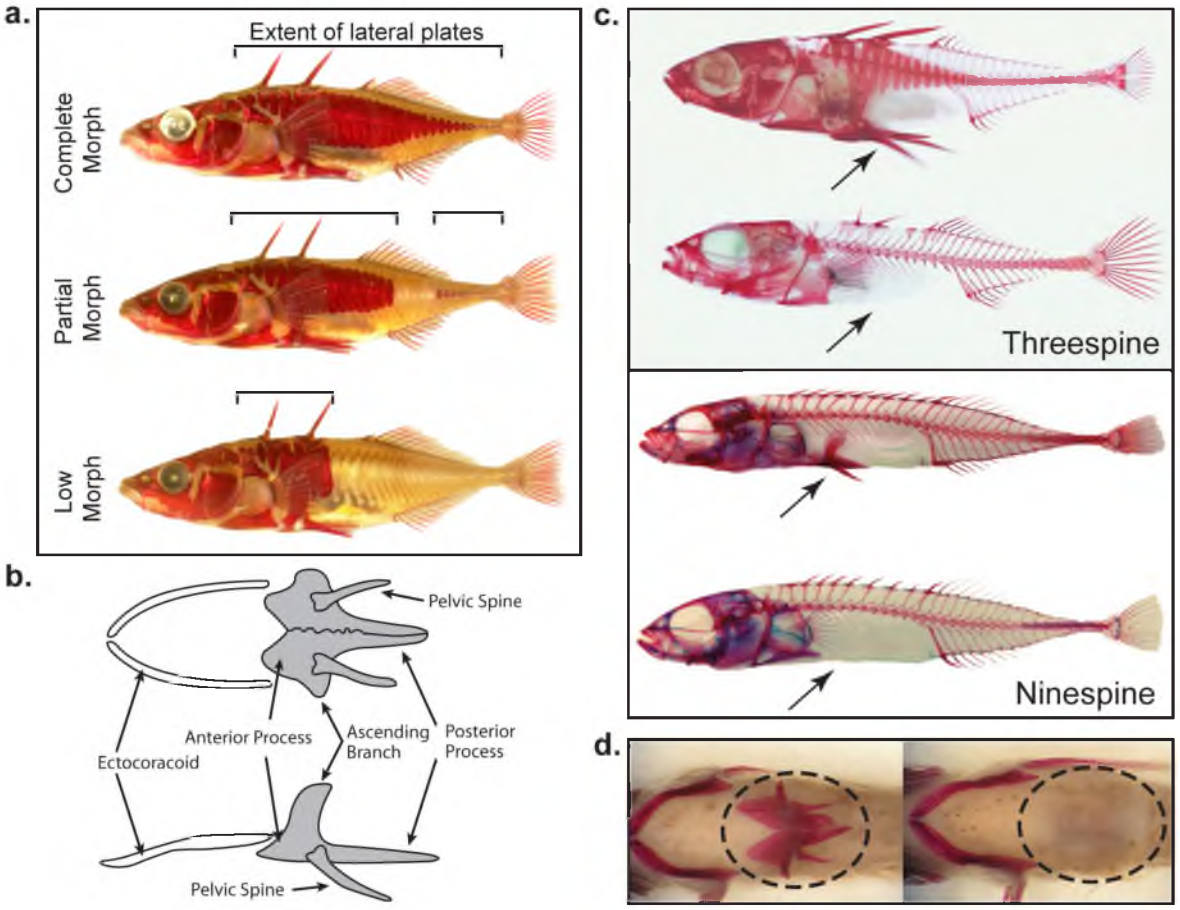
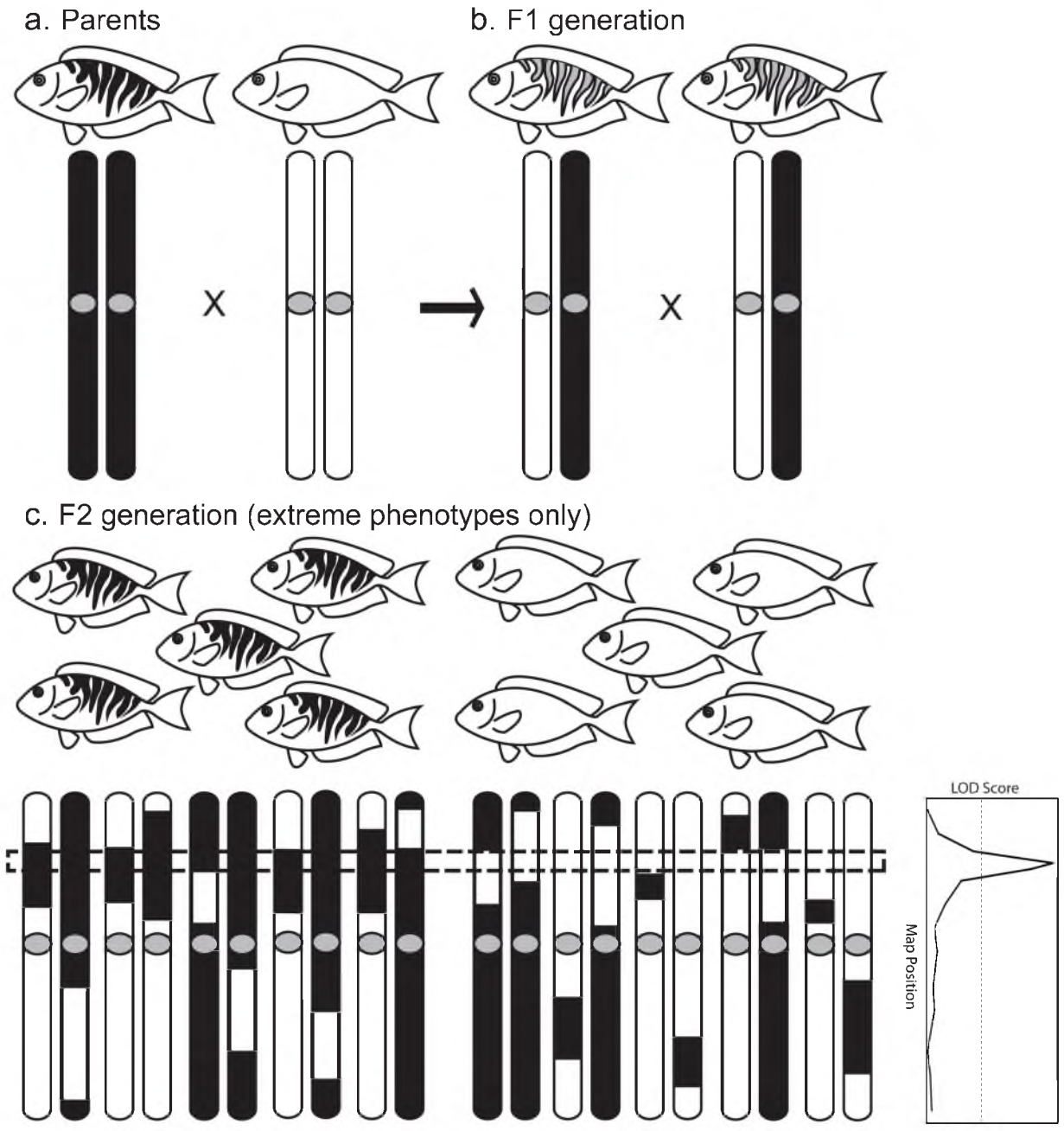


Figure 3.1. Variation in stickleback plate and pelvic phenotypes. a) Variation in lateral plate number in marine threespine sticklebacks: complete morph (top), partial morph (center), and low morph (bottom). Bony structures in all panels were visualized by staining with alizarin red. Fish found in marine habitats nearly always possess 30 or more plates per side (a phenotype referred to as the "complete morph"). In freshwater, fish typically have less than 10 plates per side ("low morph") or, less frequently, have an intermediate number of plates ("partial morph") (Hagen and Gilbertson 1972). Partial morphs exhibit a stereotypical pattern of plate loss, with plates at the most anterior and most posterior regions of the body and a midbody gap in between. Images courtesy of Jun Kitano, modified after Kitano et al. (2008). b) Ventral (top) and lateral (bottom) illustrations of the stickleback pelvis and ectocoracoid. The ectocoracoid is located anterior to the pelvis. c) Pelvic loss has evolved in multiple populations of freshwater threespine sticklebacks (*G. aculeatus*) (top) and ninespine sticklebacks (*P. pungitius*) (bottom). In both species, the ancestral marine populations possess a complete pelvis; therefore, this trait has evolved independently in each species. d) Ventral view of ninespine sticklebacks with a complete pelvis (left) and a missing pelvis (right).

Figure 3.2. Schematic of quantitative trait locus mapping in laboratory crosses. a) Individuals or populations that show variation in a trait of interest (in this case stripes) are crossed to produce F1 offspring (b), which exhibit a phenotype intermediate to the phenotypes of the parental generation. F1 individuals are crossed to produce an F2 generation (c), which will now show segregation of the trait of interest if the number of genes controlling the trait is small. In this case, only the phenotypic extremes (dark stripes or no stripes, but not intermediate stripes) are shown. The genomes of the F2 individuals are then analyzed with a set of genomic markers to detect statistical associations between genotypes and phenotypes. These associations define QTL, which are chromosome regions that are linked to phenotypes of interest. The identity of the specific genes that underlie phenotypic variation might not immediately be known because QTL associations often span many genes. Chromosome segments inherited from the striped and unstriped founders of the cross are indicated by black and white, respectively, and only the chromosome containing the causative mutation is depicted here. If individuals that inherit one version of the chromosome segment (black) nearly always exhibit one phenotype (dark stripes) and individuals inheriting the alternative version (white) nearly always exhibit the alternative phenotype (no stripes), then that segment is probably linked (physically close on a chromosome) to the causative mutation. In this example, a dashed box indicates the chromosome region associated with the stripe trait. The different versions of the chromosomes can be detected using markers such as polymorphic microsatellite markers (short repeat sequences that often differ in length among individuals) and single nucleotide polymorphisms (SNPs). These markers are assembled into linkage groups, and relative marker positions are determined based on recombination rates. Ideally, each chromosome in the genome will be represented by a single linkage group, and together the groups comprise a linkage map. Likelihood of odds (LOD) scores provide a statistical test of associations between genotypes and traits. The LOD plot at bottom right shows a region of a chromosome that exceeds a significance threshold (dashed line) and is therefore associated with variation in the trait of interest.


a. Parents
b. F1 generation
X
X
c. F2 generation (extreme phenotypes only)
LOD Score
Map Position

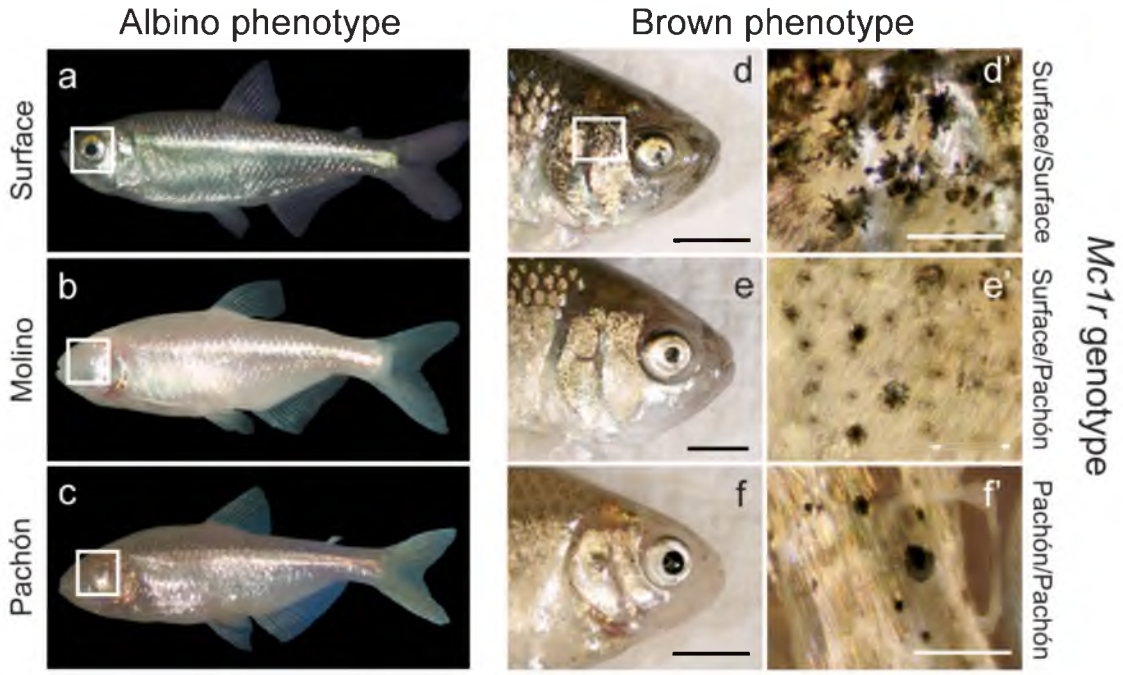


Figure 3.3. Eye loss and pigmentation differences in Mexican cavefish. Surface morph of *Astyanax mexicanus* (a) compared to cavefish populations from the Molino (b) and Pachón (c) populations. Each of these cave populations exhibits pigment loss mediated by *Oca2* and eye reduction (white boxes). In some populations, these changes have probably evolved independently. (d–f') The partially pigmented "brown" phenotype results from a decrease in melanin content and number of melanophores (pigment-containing cells). The severity of the phenotype depends on the number of cave alleles of *Mc1r* in an individual. In this example, two copies of the Pachón allele yield the most severe phenotype. Boxed area in (d) indicates area of magnification in (d'–f'). (a–c) Modified after Gross et al. (2009).

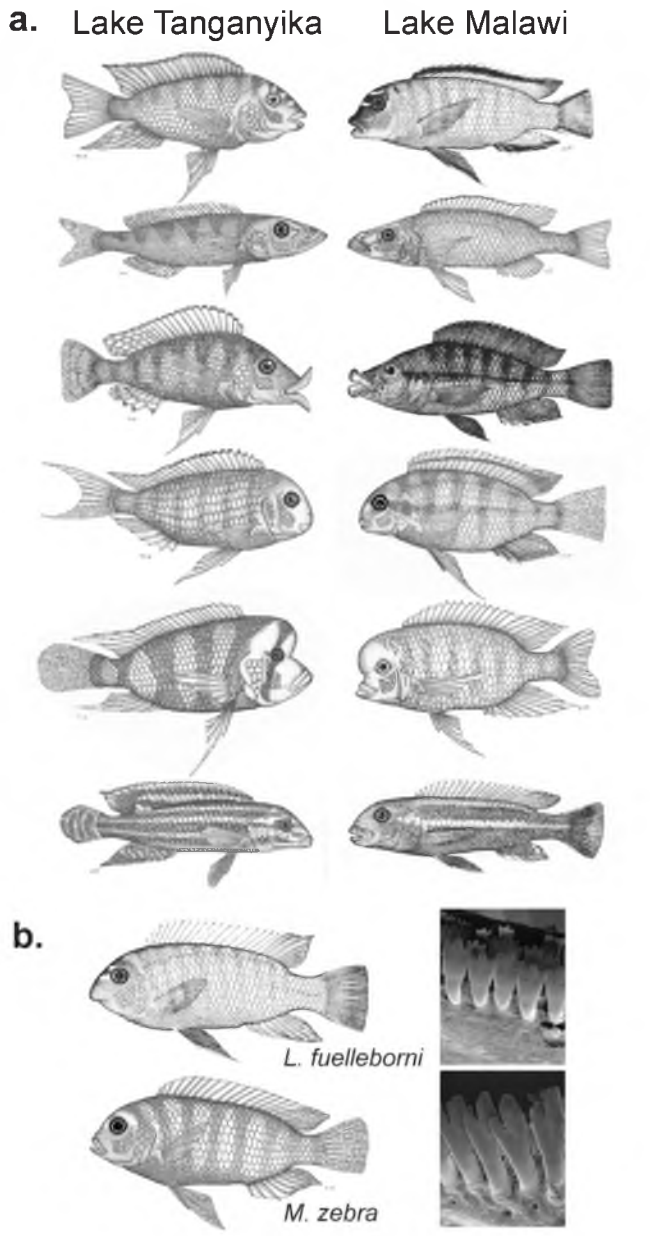


Figure 3.4. A sample of the cichlid diversity in Lake Tanganyika and Lake Malawi. a) A sample of the cichlid diversity in Lake Tanganyika (left) and Lake Malawi (right), highlighting the convergent phenotypes that have evolved independently in these two lakes. b) *Labeotropheus fuelleborni* (top) feeds by biting algae from rock surfaces. This species has a shorter lower jaw (center) and tricuspid teeth (left). In contrast, *Metriaclima zebra* (bottom) is a suction feeder with a long lower jaw and bicuspid dentition. Modified after Albertson and Kocher (2006).